STRATEGIE D'OPTIMISATION DES RECETTES DU PEAGE ROUTIER CAMEROUNAIS.

Par :

TCHEUDJEU TIEMENY Placède Judicaëlle

Maître ès Sciences

Dirigé par :

Pr. Henri GWET

Chef de Département de Mathématiques et Sciences Physiques à l'ENSP de Yaoundé.

et

M. NGAHZI Gaspard

Chef Service des Etudes à la Division des Etudes et Synthèses Générale du Budget - MINFI.

Octobre 2007

Dédicaces

A ma famille ; pour le leitmotiv que vous avez toujours été pour moi. Maman, papa, Patrick, Nadine, Christian, c'est pour vous et grâce à votre amour que je n'ai jamais baissé les bras.

Remerciements

Je commencerai par remercier le Pr. Henri GWET, le Dr Eugène-Patrice NDONG NGUEMA ainsi que tout le corps enseignant du Master II de Statistique de l'ENSP pour tout le dévouement qui a été le leur tout au long de l'année académique.

Je remercierai également messieurs Isaac JOUONANG, Samuel MEKONTCHOU, Jean-Claude TIENTCHEU, Bernard DILANGUE ainsi que tout le personnel de la Division des Etudes et Synthèses de la Direction Générale du Budget du Ministère des Finances pour l'accueil, l'espace de travail et des conseils formidables dont j'ai pu bénéficier tout au long de mon stage. Ma reconnaissance va également à l'endroit de M. Guillaume MANKOLO de la Direction Générale des Impôts, M. Patrice LODJOU de la Caisse de Stabilisation des Prix des Hydrocarbures (CSPH), M. Miché ATENGAN du Ministère des Transports et M. Donnat TAKUETE du Ministère des Travaux Publics pour tous les efforts et pour le temps qu'ils ont consacré afin que je puisse entrer en possession des données.

Je remercie les familles TIEMENY, KOUNG, DJOUMESSI, NANA, la grande famille NGANKOUE où je pense particulièrement à mes grands parents M et Mme NGANKOUE, à Guy NJAMEN et à Manuela NGANLO pour tous leurs encouragements et leur soutien permanent.

Je pense également à mes amis de toujours : Alex AKOUAGUE, Alex LONKONG, Aristide GAGNANG, Aubin NGUEYA, Carole BOUENDEU, Idosie AZANGUE, Jeannot NOUBISSIE, Nathalie CHIWA, Pamela NGUESSEU, Serge TCHANDA et Virginie NGOUTANE. Je n'oublie pas mes camarades de promotion avec qui nous avons partagé une année mémorable ; je pense particulièrement à tous les Modzans et à Cyprien MBOGNING.
Merci à tous ceux qui de près ou de loin m'ont aidé à la réalisation de ce travail.

Je terminerai par l'éternel notre Dieu sans qui rien n'est possible.

Abréviations

ACF : Autocorrelation Function
AR : Autoregressive
ARIMA : Autoregressive Integrated Movig Average
ARMA : Autoregressive Moving Average
CISOPR : Comité Interministériel de Suivi des Opérations de Péage Routier
CRPH : Caisse de Régulation des Prix des Hydrocarbures
CVS : Corrigé des Variations Saisonnières
DPO : Direction par Objectif
MA : Moving Average
MINAT : Ministère de l'Administration Territoriale
MINFI : Ministère des Finances
MINTRANS : Ministère des Transports
MINTP : Ministère des Travaux Publics
PGP : Prix du Gasoil à la Pompe
PMC : Prix Moyen du Carburant
PSP : Prix du Super à la Pompe
PSRR : Programme de Sécurisation des Recettes Routières
PPTE : Pays Pauvre Très Endetté
RUR : Redevance à l'Usage de la Route
TMI : Trafic Moyen Imposable
VPAC : Volume du Parc Automobile Camerounais

Résumé

Les sources de financement pour la construction et l'entretien des routes est un problème majeur pour les Etats de pays du tiers monde comme le nôtre. C'est ainsi que l'Etat camerounais institue le 07 Janvier 1993 le péage sur les routes bitumées afin d'assurer leur maintenance.

Le travail que nous avons effectué à la Division des Etudes et Synthèses de la Direction Générale du Budget du Ministère des Finances consistait à faire une analyse statistique des recettes du péage routier visant à trouver les moyens d'optimiser celles-ci.

Pour le faire, nous avons commencé par faire une prévision temporelle des recettes en nous servant successivement de la tendance générale et des prévisions corrigées tenant compte des variations saisonnières. Ensuite, pour inclure les résidus dans notre modélisation, nous avons standardisé cette série temporelle et avons modélisé la série résultante en un modèle $ARIMA(5, 1, 18)$. Pour améliorer la qualité de notre analyse, nous avons prix en compte d'autres variables de notre économie pouvant influencer l'évolution des recettes du péage routier et nous avons construit un modèle de prévision basé sur une régression linéaire multiple.

Ce modèle va permettre au Programme de Sécurisation des Recettes routières de :
- Mieux maîtriser le taux de fraude sur l'ensemble des postes de péage du pays ;
- Fixer un quota raisonnable pour ce qui est de la participation du péage dans la construction et l'entretien du réseau routier camerounais.

Mots clés : péage routier ; recettes ; séries temporelles ;
régression linéaire multiple.

Abstract

The sources of financing the construction and the maintenance of road is a major problem faced by third world countries in general and Cameroon in particular. It is for this reason that the state of Cameroon instituted on the 7^{th} January 1993, toll gates on tared roads to ensure their maintenance.

The study that we have carried out in the department of Studies and Compilation, of the General Directorate of Budget in the Ministry of Finances consisted of analyzing statistically the revenue from tool gate so as to look at ways and means of its optimization.
To do this , we started by carrying out a temporal forecasting of revenue using successively the general trend and adjustment forecasting of seasonal variations. After to include the residuals into our model, we standardize that temporal series and try to model the resultant series into $ARIMA(5, 1, 18)$ model. In order to improve upon the quality of our analysis we took into account other variables of our economy which could influence the evolution of the toll gate revenue from our tarred roads and we had to construct o forecasting model based on a multiple linear regression.

This model will permit the "Programme de Sécurisation des Recettes Routières" to :

- Better handle the level of fraud on the entire toll gate stations in the country ;
- Fix a reasonable quota as far as the contribution of toll gates in the construction and maintenance of road network in Cameroon.

Key words : toolgate ; revenue ; temporal series ; multiple linear regression.

Table des figures

Liste des tableaux

Introduction

Dans la civilisation des communications, les routes ont un rôle très déterminant à jouer dans le développement économique. Leur importance, tant sur le plan de la densité que de la qualité, constitue le support préalable de tout décollage économique. C'est la raison pour laquelle les pouvoirs publics, aussi bien dans les pays développés et que ceux en voie de développement, sont amenés à déployer des investissements considérables pour leur équipement en infrastructures routières.

Ainsi donc, l'Etat camerounais a jusqu'ici consenti de nombreux efforts en vue du développement de son réseau routier. Ce qui lui a permis de disposer aujourd'hui d'un réseau bitumé de plus de 4054 km contre 2796 km en 1986 [9]. Cependant, les répercussions négatives et prolongées de la crise économique sur la trésorerie de l'Etat ont considérablement bouleversé cet important programme routier. L'Etat Camerounais, principal garant du développement et de l'entretien du réseau bitumé depuis son accession à l'indépendance, n'arrive plus à faire face à ses engagements en cette matière.

Il était donc question de définir une nouvelle politique en matière d'infrastructures routières. L'impératif du développement économique exige qu'un accent particulier soit mis sur l'élaboration et la mise en application d'une politique routière qui tienne compte des réalités de l'heure, à savoir : rétrécissement et rareté des financements du réseau routier, entretien et extension du réseau bitumé, désenclavement du pays.

C'est dans cette optique que l'Etat a procédé à l'élargissement de son assiette fiscale en instituant une nouvelle taxe : le péage routier. Pour lui, cette taxe devait permettre le recouvrement des charges d'entretien dues à l'usage du réseau bitumé par les automobiles.

A sa quatorzième année d'expérience, nombreux sont ceux qui s'interrogent sur les performances du péage. En d'autres termes, le péage a-t-il permis d'atteindre les objectifs fixés lors de son instauration ? Existe-t-il des voies et moyens pouvant permettre de corriger les imperfections du péage routier actuellement en vigueur et de renouer

avec les impératifs d'efficacité et de rentabilité ? Notre étude à la Division des Etudes et Synthèses de la Direction Générale du Budget du Ministère des Finances a eu pour but de trouver des stratégies pour une optimisation réelle des recettes du péage routier camerounais.

Pour le faire, nous avons articulé notre travail en quatre chapitres :

- Dans le premier, nous faisons un bilan du péage routier camerounais.
- Le deuxième chapitre est une présentation et une description des données à notre disposition.
- Le troisième chapitre est une présentation des méthodes statistiques utilisées dans la réalisation de notre étude.
- Le quatrième chapitre consiste en l'application des différentes méthodes exposées au troisième chapitre aux données et à la présentation des résultats obtenus.

Résumé exécutif

Sujet

Stratégie d'optimisation des recettes du péage routier camerounais.

Problème

Du fait de la rareté de fonds de financement pour la construction et l'entretien du réseau routier, il a été institué au Cameroun le 07 janvier 1993 le péage routier. Du fait de la non atteinte des objectifs qui lui étaient assignés au départ, le PSRR (Programme de Sécurisation des Recettes Routières) est mis sur pied en Octobre 2005, avec pour objectif (comme son nom l'indique) de sécuriser les recettes issues de la route qui sont : la RUR (Redevance à l'Usage de la Route), la taxe à l'essieu, les amendes routières et le péage routier.

Le péage se devait de fournir 5.7 milliards de FCFA en 2006 (ce qui n'a pas été le cas) en vue d'atteindre 08 milliards de FCFA en 2012, raison pour laquelle il est urgent de trouver des solutions visant à améliorer son efficacité.

Actuellement, le PSRR fixe un taux appelé DPO (Direction Par Objectif) à chaque poste de péage de la façon suivante : en fonction du poids du poste dans les recettes des années précédentes et de l'accroissement du Volume du Parc Automobile Camerounais (VPAC) - il est supposé chaque année une augmentation du volume du parc automobile camerounais de 7% et un retrait de la circulation de 3% des véhicules - il est fixé pour chaque poste un montant mensuel à atteindre. Si ce montant est dépassé, la prochaine DPO, du poste sera diminuée du surplus et s'il n'est pas atteint, il sera augmenté dans la prochaine DPO la somme restante. Cette façon informelle de prévoir les recettes du péage se doit d'être améliorée ; raison pour laquelle il est urgent de mener une analyse statistique d'optimisation des recettes du péage qui prendra en compte d'autres paramètres incontournables comme, par exemple, le Trafic Moyen Imposable (TMI) sur l'ensemble des postes de péage du pays.

Données

Pour réaliser la prévision qui est le but de notre étude, nous allons chercher à modéliser la chronique des recettes mensuelles du péage routier allant de Juillet 1995 à Décembre 2005. Nous avons également à notre disposition les données sur le trafic moyen imposable sur l'ensembles des postes de péage du pays, le volume du parc automobile camerounais, le prix du super à la pompe et le prix du gasoil à la pompe.

Résultats

Prévision temporelles

Tendance générale

De 1995 à 2005, les recettes suivent une tendance essentiellement linéaire donnée par la droite d'équation (donnée par la méthode des moindres carrés) :

$$\widehat{Y}_t = 2023667\,t + 204417147, \tag{1}$$

où t représente le numéro du mois dans l'année ; le mois N° 1 étant juillet 1995.

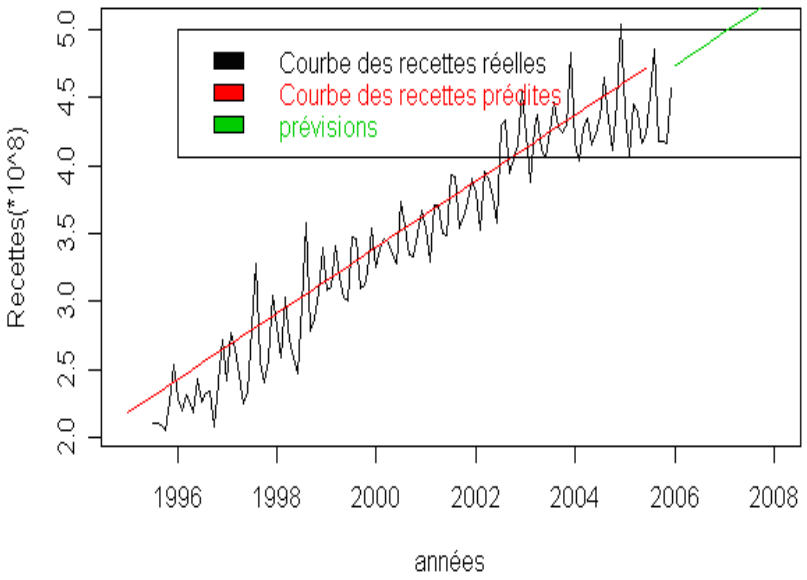


FIG. 1 – Tendance des recettes du peage routier Camerounais.

A base de cette tendance générale, voici les résultats qu'on aurait pu avoir pour les prévisions de janvier 2006 à décembre 2008.

Mois	2,50%	Prévisions	97,50%
janv-06	426012839	473564866	521116893
févr-06	428018820	475588533	523158246
mars-06	430024532	477612200	525199869
avr-06	432029974	479635867	527241761
mai-06	434035147	481659534	529283922
juin-06	436040051	483683202	531326352
juil-06	438044687	485706869	533369050
août-06	440049055	487730536	535412016
sept-06	442053155	489754203	537455250
oct-06	444056988	491777870	539498752
nov-06	446060554	493801537	541542520
déc-06	448063853	495825204	543586555
janv-07	450066885	497848871	545630857
févr-07	452069652	499872538	547675424
mars-07	454072153	501896205	549720257
avr-07	456074388	503919872	551765356
mai-07	458076359	505943539	553810719
juin-07	460078065	507967206	555856348
juil-07	462079506	509990873	557902240
août-07	464080684	512014540	559948396
sept-07	466081599	514038207	561994816
oct-07	468082250	516061874	564041499
nov-07	470082638	518085542	566088445
déc-07	472082764	520109209	568135653
janv-08	474082628	522132876	570183123
févr-08	476082230	524156543	572230855
mars-08	478081572	526180210	574278848
avr-08	480080652	528203877	576327102
mai-08	482079472	530227544	573375616
juin-08	484078032	532551211	580424390
août-08	486076332	534274878	582473424
août-08	488074373	536298545	584522717
sept-08	490072155	538322212	586572269
oct-08	492069679	540345879	588622079
nov-08	494066945	542369546	590672147
déc-08	496063954	544393213	592722473

TAB. 1 – Prévisions sur les recettes du péage de 2006 à 2008 à base du modèle (1).

Prévision corrigée : en tenant compte des variations saisonnières

Dans notre étude, nous avons montré qu'il existe un mouvement saisonnier assez clairement identifiable dans les recettes du péage routier camerounais. En tenant compte de ce mouvement saisonnier, on peut faire des prévisions à priori plus précises de l'évolution de ces recettes dans le temps en s'appuyant cette fois sur le modèle défini par l'équation (2) suivant qui corrige le modèle défini par l'équation (1) :

$$\widehat{Y}_t = 2023667\,t + 204417147 + S_t, \tag{2}$$

où S_t représente le coefficient saisonnier du mois dans l'année calendaire. Ces coefficients saisonniers sont donnés dans le tableau 2 : Des prévisions plus précises pour les recettes

Mois	S_t
Janvier	13359120
Février	27731668.7
Mars	-11075940.4
Avril	-17355574.4
Mai	-1730671.6
Juin	32577011.2
Juillet	-1353520
Août	-14890682
Septembre	9899669.6
Octobre	474661.2
Novembre	-18029449.2
Décembre	-19606292.9

TAB. 2 – Coefficients saisonniers des différents mois.

du péage routier de janvier 2006 à décembre 2008 basées sur (2) sont fournies dans le tableau 3 :

La figure 2 présente la série corrigée des variations saisonnières, ainsi qu'une estimation de la tendance globale à long terme de la chronique des recettes.

Prévision en fonction d'autres variables économétriques pertinentes

Deuxièmement, dans notre étude nous avons essayé de voir s'il était possible d'introduire d'autres variables de notre économie qui permettraient de mieux expliquer l'évolution des recettes ; Nous avons ainsi identifié les variables :

- TMI : Trafic Moyen Imposable ;
- VPAC : Volume du Parc Automobile Camerounais ;
- PSP : Prix du super à la Pompe ;

Mois	2,50%	Prévisions	97,50%
janv-06	404112926	474138135	544163345
févr-06	406114138	476165393	546216647
mars-06	408114954	478192650	548270346
avr-06	410115373	480219908	550324442
mai-06	412115396	4822447165	552378933
juin-06	414115024	484274422	554433821
juil-06	416114256	486301680	556489103
août-06	418113093	488328937	558544781
sept-06	420111537	490356195	560600852
oct-06	422109586	492383452	562657318
nov-06	424107242	494410709	564714176
déc-06	426104505	496437967	566771428
janv-07	428101377	498465224	568829072
févr-07	430097856	500492481	570887107
mars-07	432093944	502519739	572945534
avr-07	434089641	504546996	575004351
mai-07	436084948	506574254	577063559
juin-07	438079866	508601511	579123156
juil-07	440074394	510628768	581183143
août-07	442068534	512656026	583243518
sept-07	444062285	514683283	585304281
oct-07	446055649	516710541	587365432
nov-07	448048626	518737798	589426970
déc-07	450041217	520765055	591488894
janv-08	452033422	522792313	593551204
févr-08	454025242	524819570	595613899
mars-08	456016677	526846828	597676978
avr-08	458007728	528874085	599740442
mai-08	459998395	530901342	601804290
juin-08	461988680	532928600	603868520
août-08	463978582	534955857	605933132
août-08	465968103	536983115	607998127
sept-08	467957242	539010372	610063502
oct-08	469946001	541037629	612129258
nov-08	471934380	543034887	614195394
déc-08	473922380	545092144	616261908

TAB. 3 – Coefficients saisonniers des différents mois.

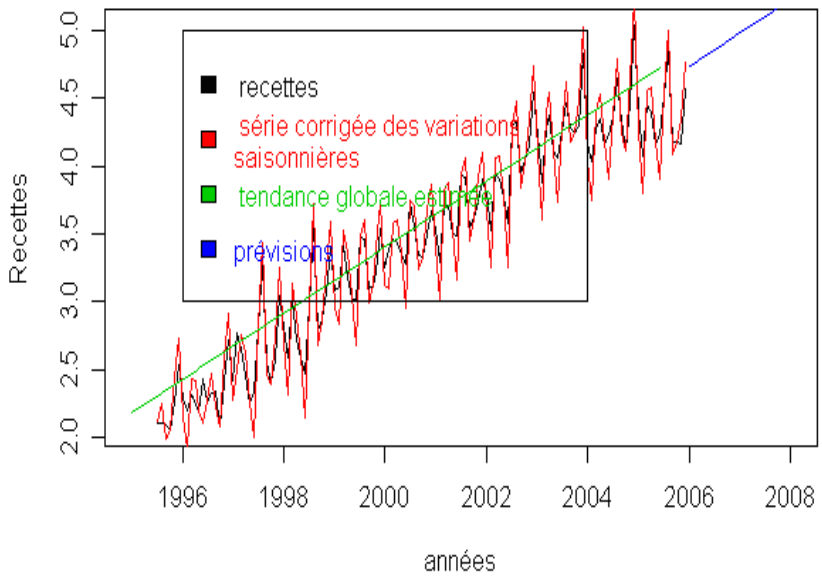


FIG. 2 – Graphe de la série corrigée des variations saisonnières.

– Prix du Gasoil à la Pompe ;
– Prix Moyen du Carburant.

Ceci nous a permis de réaliser le modèle économétrique suivant permettant d'expliquer et de prévoir les recettes une fois connues les réalisations des variables ci-dessus.

$$\widehat{Y} = TME^{1.07} \times VPA^{0.75} \times PSP^{1.52} \times PGP^{-0.94}.$$

Ce deuxième modèle peut avoir des intérêts de prédiction à très court terme (échéance 1 mois) des recettes mensuelles du péage. En effet les différentes variables explicatives de notre modèle s'observent chaque mois avant les recettes : le trafic moyen imposable et les prix des carburants pour un mois quelconque sont connus au plus tard au début de ce mois, les données sur le volume du parc automobile quant à elles sont annuelles alors que les données sur les recettes sont disponibles au plus tôt le 1^{er} du mois suivant.

Erreurs quadratiques d'ajustement

Nous présentons dans le tableau 4, les erreurs quadratiques liées aux différents ajustements.

	Nombre de mois	er(%)	em
Tendance générale	**132**	**6.63**	**22920236**
Prévisions corrigées	**132**	**5.76**	**13752385**
Modèle économétrique	**132**	3.01×10^{-8}	**0.1042857**

TAB. 4 – Erreurs quadratiques d'ajustement.

Suggestions et recommandations

De cette étude, nous pouvons suggérer des réformes à appliquer au système de péage actuel et des innovations à réaliser dans le futur. C'est la raison pour laquelle ces suggestions portent sur le court ou moyen terme et sur le moyen ou long terme.

Réformes à court ou moyen terme

Elles portent sur :

- **La réforme de la fiscalité routière** : l'analyse du niveau des contributions fiscales des usagers de la route a montré que certaines catégories d'usagers devraient payer plus que d'autres pour compenser les dégâts qu'ils occasionneraient au réseau. Il faudrait donc mettre un accent particulier sur l'allègement du fardeau fiscal des véhicules de transport des voyageurs et des utilitaires légers. On évitera, en outre, de les frapper par des nouvelles taxes fiscales dans le but d'encourager leur contribution au péage routier. Car le relèvement du péage dépend énormément de ces deux catégories d'usagers qui représentent, à eux seuls, sensiblement 60% du parc automobile national. Il convient aussi de souligner le caractère nocif de la multiplicité des taxes fiscales qui frappent l'ensemble des usagers de la route. Certaines d'entre elles comme la vignette, la taxe à l'essieu et la taxe de dégradation de la chaussée font double emploi. Elles devraient par conséquent être fondues en une seule. Il serait opportun d'instaurer un guichet unique qui aurait le triple avantage de :
  - Limiter la fraude ;
  - Simplifier la vie des contribuables ;
  - Faciliter la tâche de l'administration.
- **La lutte contre le péage informel** : il a été révélé dans [9] qu'il existe une corrélation négative entre le nombre de postes de contrôle des forces de l'ordre sur les routes nationales et le niveau de contribution des usagers au péage. Il est donc impératif de réduire ces contrôles qui enrichissent considérablement le poids de la fiscalité pour les usagers de la route en général et des routes interurbaines en particulier. Cette mesure aura sans doute une incidence positive sur la rentabilité du péage.

- **La réfection du réseau bitumé interurbain en vue de son adaptation aux normes du péage routier** : le réseau interurbain camerounais présente des caractéristiques impropres conventionnelles du péage. En plus du mauvais état physique de certaines routes, des problèmes liés à la maintenance du réseau, de l'absence des itinéraires alternatifs, le réseau est 1×2 voies alors que la norme requise est de 2×2 voies au minimum. Compte tenu de cette situation, les pouvoirs publics, dans l'optique de la restauration du péage routier, devront donc, dans un proche avenir, définir et mettre en œuvre un programme de réfection des routes à péage. L'objectif visé consistera à mettre à la disposition des usagers un réseau offrant des garanties de sécurité, de confort et leur permettant de réaliser des gains de temps substantiels.

Réformes à moyen ou long terme

Nous pouvons identifier trois réformes essentielles à entreprendre :

- **L'automatisation des postes de péage** : la fraude au péage est due en majorité au fait que la perception du droit de péage se fait manuellement. Aucun poste sur le réseau bitumé camerounais ne dispose d'un compteur automatique de véhicules. La perception par les équipements automatiques tels que les récepteurs de pièces de monnaie encore appelés « paniers » et les lecteurs de cartes magnétiques est la plus utilisée dans les pays développés. L'usage de ces équipements dans la perception des droits du péage routier présente un grand nombre d'avantages. On note une plus grande productivité des équipements du fait de la rapidité de la perception, donc de l'écoulement du trafic. En plus, ces équipements de perception sont accompagnés de dispositifs qui permettent d'éviter la fraude, de suivre les opérations de façon détaillée et de recueillir des statistiques. Il faut également souligner la garantie de la sécurité des fonds surtout avec l'usage de la monétique (cartes magnétiques et bancaires pour péage routier). Car dans le cas de cette technique de perception, le paiement du droit de péage a lieu à la source et non pas au poste de péage.

- **La redynamisation du programme de sécurisation des recettes routières (PSRR)** : Afin de garantir le meilleur usage des recettes provenant du péage et de procéder à un meilleur entretien des routes et à la construction de quelques autres. Ce programme se doit de mieux canaliser l'aide internationale et toutes les ressources issue du fonds routier parmi lesquelles le péage ; il devrait permettre de faire la lumière sur l'argent du péage, car ce dernier n'a pas servi à ce à quoi il était destiné. La redynamisation du PSRR est un passage obligé car en raison de son autonomie financière, elle ouvre la voie à la consécration des recettes du péage aux seules fins de maintenance du réseau et de fonctionnement

de l'appareil de gestion du péage routier. Il s'agit de gérer les routes comme une entreprise au lieu de les administrer comme un service social.

- **La privatisation de la gestion du péage routier** : face aux échecs qu'a occasionné la gestion en régie de l'infrastructure routière, un certain nombre d'Etats ont adopté l'option de la privatisation. La France est par exemple l'un des pays où cette stratégie moderne a été mise en œuvre. Les résultats obtenus jusqu'à ce jour sont très satisfaisants. Cette privatisation apparaît comme la seule voie de salut susceptible d'améliorer l'état de nos réseaux routiers.

Conclusion Générale

Cette étude nous a permis d'identifier les causes réelles de l'échec du système de péage routier actuel, de formuler des propositions de réformes et de modéliser afin de prévoir les recettes à venir du péage routier pour les exercices budgétaires à venir.

Pour réaliser nos prévisions, nous avons premièrement abordé le problème à l'aide de prévisions temporelles. Ensuite nous avons associé aux recettes différentes variables de notre économie ayant un impact sur le péage à base desquelles il nous a été possible de construire un modèle basé sur la régression linéaire multiple. Le premier modèle a des intérêts de prévision à moyen ou à long terme alors que le second est utile pour des prévisions à très court terme (échéance 1 mois) des recettes mensuelles du péage.

Nous ne saurons terminer sans noter le fait qu'une approche basée sur les séries temporelles multivarieés est un chemin qu'il reste à explorer pour ce qui est des prévisions des recettes du péage routier camerounais.

MODE D'EMPLOI DU PROGRAMME DE PREVISION DES RECETTES DU PEAGE ROUTIER CAMEROUNAIS

Téléchargement et installation du logiciel R

Le logiciel R est une version non commerciale de S-plus. Télécharger ce logiciel sur le site Internet **http ://lib.stat.cmu.edu/R/CRAN/** ou sur le site **http ://cran.ch.r-project.org/**, installer ensuite ce logiciel dans un ordinateur.

Création d'un répertoire de travail

a- Créer un dossier de travail dans l'ordinateur où est installé le logiciel R. On pourra, par, exemple créer ce dossier dans la racine « C » et le nommer « **outils pour la prévision des recettes** » ;

b- Copier dans notre dossier de travail, les programmes de prédiction « C.40 » et « C.54 » se trouvant en Annexe du mémoire. Enregistrer ensuite ces fichiers dans le dossier de travail précédemment créé en les nommant par exemple « **Rec.pred1** » et « **Rec.pred2** » ;

c- On suppose maintenant que les données mensuelles (Recettes, TMI, VPAC, PSP, PGP, PMC) qui ont servi à l'écriture de nos modèles sont enregistrées dans notre dossier de travail sous la forme d'un tableau Excel nommé, par exemple, « **données mensuelles du péage** ».

Importation du programme de prédiction et du tableau des données dans la console de R.

a- Lancer le logiciel R en double cliquant par exemple sur son icône ;

b- Changer le répertoire courant de R au dossier de travail précédemment créé en exécutant la commande suivante :
setwd("indiquer le chemin d'accès au dossier de travail créé"),
par exemple,
setwd("C :/outils pour la prévision des recettes")
dans le cas où les noms proposés ont été utilisés ; puis, valider cette instruction en cliquant sur la touche du clavier « **Entrée** » ;

c- Importer le programme nécessaire à notre prévision dans la console de R en exécutant la commande suivante :
source(" nom du fichier contenant le programme de prédiction ")
qui est par exemple ceci :
source("Rec.pred1")
dans le cas où le nom du fichier contenant ce programme de prédiction proposé a été utilisé ; puis, valider cette instruction en cliquant sur la touche du clavier « **Entrée** » .

Exécution de la prévision des recettes mensuelles du péage.

Effectuer la prévision en utilisant la commande suivante :
predict (reg2,x)
où x est le tableau contenant les numéros des mois où on voudrait faire une prévision des recettes ; dans le cas des prévisions à moyen ou long terme,
predict(REG(tmi,vpac,psp,pgp))
où les variables tmi,vpac, psp et pgp désignent respectivement le trafic moyen imposable, le volume du parc automobile camerounais, le prix du super à la pompe et le prix du gasoil à la pompe correspondant au mois où on voudrait faire une prévision des recettes ; dans le cas des prévisions à échéance d'un mois.

Chapitre 1

Bilan et problématique du peage routier Camerounais

1.1 Historique

L'idée du péage routier a été émise au Cameroun pour la première fois en 1977. Ce fut à l'issue de l'étude de création d'un fond routier réalisée par une firme italienne pour le compte du Ministère de l'Equipement. Mais cette idée n'a pas été creusée.

Lors de la session parlementaire de juin 1986, les députés avaient noté avec satisfaction la construction de l'axe lourd Yaoundé-Douala. C'est alors qu'ils avaient émis le souhait de voir instituer le péage sur les routes bitumées pour assurer leur maintenance.

Le Ministère des Transports a, à cet effet, préparé un projet d'ordonnance qui a été soumis à la signature du chef de l'Etat en 1988. En même temps, l'étude menée par la Direction des Transports recommandait la baisse du prix du carburant comme mesure d'accompagnement de l'institution du péage. Aucun acte n'ayant été pris en faveur de la baisse du prix du carburant, le projet d'ordonnance suscité n'a pas connu de suite favorable.

La relance de l'institution d'un péage routier a refait surface en 1991. Elle est due à plusieurs raisons :

- la baisse du prix du carburant intervenue en mai 1991 ;
- la signature des contrats de performance avec les sociétés sous tutelle du Ministère des Transports. L'objectif de cet acte était d'aboutir à l'égalisation des conditions de concurrence entre tous les modes de transport ;
- le refus de plus en plus prononcé des bailleurs de fonds de supporter les coûts d'entretien et de renouvellement des infrastructures routières.

Aussi, la loi des Finances de l'exercice 1992/1993 institue le péage routier qui a pour objectifs :

- d'assurer le recouvrement des charges d'entretien et de renouvellement des infrastructures routières sur les usagers effectifs de la route ;

- d'accroître les ressources budgétaires de l'Etat ;
- d'égaliser les conditions de concurrence entre les différents modes de transport.

Ainsi, le décret n°93/034 du 07 janvier 1993 du Premier Ministre, chef du Gouvernement, fixe les modalités du péage sur certains axes bitumés du réseau routier national :

- le franchissement de tout poste de contrôle de péage est subordonné à la présentation d'un ticket de la valeur de 500 FCFA émis par le Ministère chargé des Finances ;
- les formules d'abonnement à tarif réduit, pour un itinéraire n'allant pas au-delà d'un poste de contrôle de péage, peuvent être consenties à des personnes physiques ou morales ayant leur domicile ou leur lieu de travail au voisinage d'un axe bitumé à péage.
- Sont exempts du droit de péage :
  - les piétons ;
  - les engins à deux roues ;
  - les ambulances ;
  - les véhicules concourant au maintien de l'ordre.

L'arrêté n°003/A/MINAT du 18 février 1993 fixe les routes à péage et crée 35 postes de contrôle (tableau 1.1 **Source** : Direction des Infrastructures et Investissements routiers (Ministère des Transports)), qui deviennent opérationnels à partir du 30 Novembre 1993.

Ce nombre est passé de 35 au départ à 45 aujourd'hui en raison de la construction de nouveaux axes routiers.

1.2 Bilan financier

1.2.1 Evolution progressive des recettes

D'une manière générale, on note une tendance à la hausse des recettes du péage routier depuis sa mise en place, lesquelles sont passées de moins d'un milliard de F.CFA en 1993-1994 à plus de quatre (04) milliards de F.CFA en 2006, mais on ne peut s'empêcher de remarquer une baisse de plus de 10% entre 2005 et 2006 (figure 1.2).

Pour les douze derniers exercices budgétaires, les recettes ont évoluées ainsi qu'il suit (Tableau 1.2 **Source** : Direction générale des Impôts :Ministère des Finances) :

Au total, plus de 57 milliards de F.CFA collectés sur 14 ans d'existence du péage. Malgré la tendance à la hausse d'une année budgétaire à l'autre depuis son instauration, le péage routier n'a pas encore atteint son seuil potentiel. Un taux de fraude très élevé et une gestion contestable justifient ce manque à gagner dans les recettes.

Axes routiers	**Nombre de postes de contrôle**
Yaoundé – Bafoussam - Foumban	4
Douala - MBouda	4
Yaoundé - Douala	3
Maroua - Kousseri	3
Yaoundé - Ebolowa	2
Douala - Kumba	2
Tibati - Meidougou	2
Ngaoundéré - Garoua	2
Garoua - Maroua	2
Maroua - Yagoua	2
Edéa - Kribi	1
Mbalmayo - Sangmelima	1
Boucle de Sangmelima	1
Limbé - Idenau	1
Bafoussam - Dschang	1
Bangangté - Bafang	1
Maroua - Mokolo	1
Guider - Dourbeye	1
Bertoua - Bélabo	1
TOTAL	**35**

TAB. 1.1 – Axes routiers à péage et nombre de postes de contrôle.

Face à l'écart négatif extrêmement important entre les réalisations sur le terrain et le rendement potentiel des postes de péage, le Ministère de l'Economie et des Finances a adopté et mis en application deux importantes mesures en 1994. La première fixait, à chaque poste, un potentiel de rentabilité moyen inférieur au potentiel probable et supérieur au potentiel actuel. Quand à la seconde mesure, elle exposait les sanctions encourues par les agents d'astreinte qui n'auront pas atteint le seuil fixé à leur poste de péage. Ce sont ces mesures qui ont eu un impact positif sur le rendement global du péage entre 1995/1996 et 2005 ; mais le passage de la gestion du péage routier de la Direction Générale du Budget à celles des Impôts et le déplacement de certains postes de péage ont vu une diminution des recettes du péage routier.

On observe entre 1995/1996 et 2005 une augmentation d'un exercice budgétaire à l'autre avec un taux d'accroissement moyen d'environ 8% par an. Mais il y a en 2006 une diminution de plus de 10% ; c'est la raison pour laquelle nous nous baserons sur les recettes allant jusqu'en 2005 pour effectuer notre étude. Cependant, bien qu'alors

EXERCICES BUDGETAIRES	RECETTES FCFA
1995/1996	2 684 372 000
1996/1997	2 900 863 500
1997/1998	3 289 382 000
1998/1999	3 751 103 000
1999/2000	4 005 671 500
2000/2001	4 235 895 000
2001/2002	4 517 131 000
2002	2 530 090 500
20003	5 129 991 000
2004	5 216 479 000
2005	5 216 834 500
2006	4 670 424 275
2007	2 340 976 500

TAB. 1.2 – Evolution annuelle des recettes du péage routier.

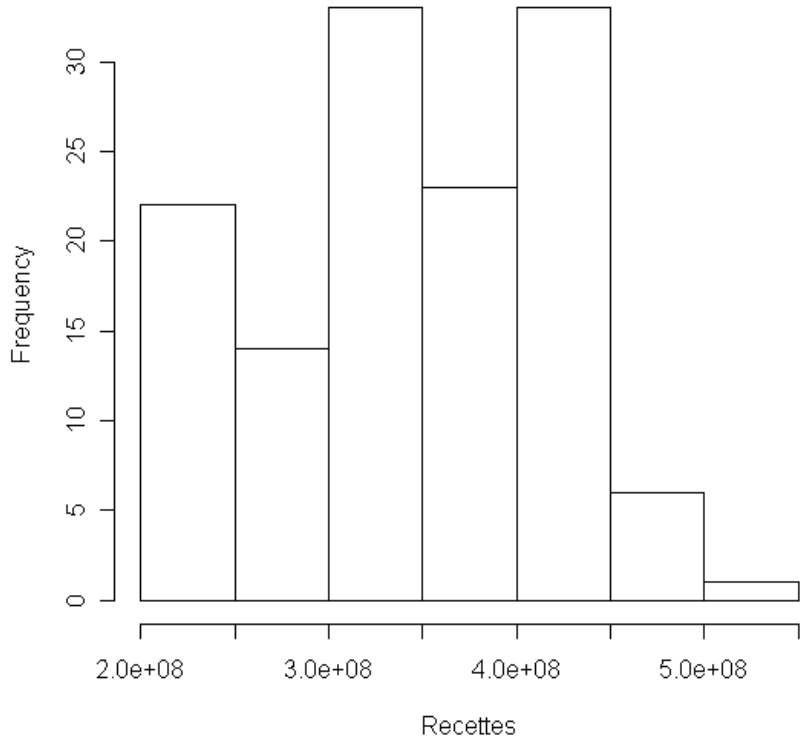


FIG. 1.1 – Histogramme des recettes du peage Camerounais.

il y ait eût accroissement, ces recettes étaient en deçà de ce qui devait être. Les causes liées à la faiblesse de ce rendement sont au nombre de trois :

- un taux de fraude élevé ;

- un coût de fonctionnement important ;
- le manque à payer des contribuables.

1.2.2 La fraude au péage

C'est l'un des problèmes majeurs rencontrés au niveau du péage. Il est parfaitement organisé, puisqu'il est impossible de suivre précisément les recettes au jour le jour.

Les fraudeurs utilisent trois techniques pour détourner les recettes. La première technique consiste à vendre des souches à la place des tickets. La deuxième s'illustre par l'usage, dans la même journée, de plusieurs carnets à souches. Quant à la troisième technique, elle se caractérise de la manière suivante : les agents de contrôle, lorsque l'occasion se présente, laissent passer deux fois le même véhicule sur la base du même ticket payé à l'aller du trajet et remis à ceux-ci au retour. Ce ticket est revendu à un autre usager par la suite. Plus fréquemment, le passage de certains véhicules se fait moyennant un cadeau de 200 à 300 FCFA, certains usagers fréquents payant encore moins.

Le contrôle réalisé en 1999 dans le cadre de l'Etude de simplification et d'harmonisation de la fiscalité routière au Cameroun au voisinage de 4 postes de péage et par interpolation linéaire sur l'ensemble des postes en activité a révélé un taux de fraude de 50%. Parallèlement, une enquête menée par un comité interministériel placé sous la direction du Ministère des Transports a enregistré pour une journée sur l'ensemble des postes :

Trafic moyen imposable	31.137 véhicules
Recette prévisionnelle	15.568.500 FCFA
Recette effective	moins de 4.000.000 FCFA
Déficit	11 millions CFA

Sur cette base, le taux de fraude estimé était de l'ordre de 73%.

1.2.3 l'importance du coût de fonctionnement de péage

L'évaluation de ce coût a été réalisée sur la base de l'Etude de simplification et d'harmonisation de la fiscalité routière au Cameroun. Les bases de cette évaluation sont les suivantes :

- nombre de postes de péage en activité : 35 ;
- nombre d'équipes par poste du péage : 3 ;
- nombre d'agents par équipe : 6 ;
- salaire mensuel par agent : 40.000 CFA ;

- charges sociales et diverses : 60

D'où un coût de fonctionnement annuel estimé à 480 millions de FCFA. La recette prévisionnelle annuelle est estimée à 4.590 millions de FCFA. Elle est déterminée sur la base d'un trafic moyen journalier de 875 véhicules par poste de péage. Cette recette se divise ainsi qu'il suit :

- part de l'Etat : 30% ;
- part personnelle des agents : 18% ;
- part des fraudeurs : 18%.

Rubrique	**Montant (millions de FCFA)**
Rentrées de l'Etat	1.377
Coût de fonctionnement	480
Part personnelle des agents (1)	550,8
Manque à payer des contribuables (2)	830
Montant de la fraude (1) + (2)	1.380,8

TAB. 1.3 – Répartition de la recette du péage.

On en déduit d'après le Tableau 1.3 (**Source** : Etude de simplification d'harmonisation de la fiscalité routière au Cameroun,1999.), un ratio avantages/coûts de l'ordre de 2,87, synonyme que les recettes effectives représentent pratiquement le triple des charges de fonctionnement. On se rend en effet compte que le péage est rentable. Malgré la mauvaise gestion due à la fraude généralisée, le rendement est encourageant même s'il demeure à un niveau insatisfaisant. Le coût de fonctionnement par ailleurs représente 35% des entrées brutes de l'Etat.

1.3 Bilan organisationnel et fonctionnel

1.3.1 Bilan organisationnel

Le péage est géré par un Comité Interministériel de Suivi des Opérations du péage routier (CISOPR) composé de :

- □ Un représentant du Ministère des Transports ;
- □ Un représentant du Ministère des Travaux publics ;
- □ Un représentant du ministère de l'Administration Territoriale ;
- □ Un représentant du Ministère de la Défense ;
- □ Un représentant du Trésor pour le Ministère des Finances ;
- □ Il est coiffé par le Directeur Général du Budget.

On note un manque de coordination entre les différentes administrations impliquées dans le péage ; nous pouvons ici penser à l'exclusion des Ministères de Travaux

Publics et des Transports qui participent pourtant considérablement à la vie de la route. D'ailleurs, c'est le Ministère des Transports qui est responsable de cette fiscalité routière. Dans une optique de maximisation des recettes, il aurait pu jouer un rôle considérable au sujet de la redynamisation des contributions des usagers de la route. Quant au ministère des Travaux publics, il est le responsable de l'entretien et de la réhabilitation du réseau routier.

En Octobre 2005 est mis sur pied le Programme de Sécurisation des Recettes Routières (PSRR) qui a pour but de gérer toutes les ressources du Fonds Routier que sont : la RUR (Redevance à l'Usage de la Route), le péage routier, les amendes routières et la taxe à l'essieu.

L'organisation du péage souffre aujourd'hui de :

- La non informatisation du système de gestion : son existence aurait certainement permis un minimum de contrôle et d'évaluation du rendement du péage.

- Manque de structure dont le rôle serait de recueillir la part des recettes destinées à l'entretien du réseau routier : du fait de l'unicité des caisses de l'Etat, tout va au trésor public.

- Gestion irresponsable et détournement des fonds des agents de contrôle au niveau local : l'évaluation des recettes détournées se situerait autour de 550,8 millions de FCFA/an et représenterait environ 40% du manque à gagner de l'Etat [8].

- L'absence de pénalités à l'encontre des contrevenants : il est difficile au regard de la formule actuelle du péage, d'instituer un cadre pénal efficace, car la gestion du péage est non seulement manuelle, mais son système est du type ouvert. Le comptage des véhicules ayant franchi un poste donné au cours d'une journée n'est pas réalisé de façon automatique. Bien plus, on n'a pas l'information précise sur la provenance de l'usager. C'est l'une des causes de la malversation financière que connaît le péage.

- L'irrationalité de la tarification en vigueur et son impact sur la rentabilité : à la traversée d'un poste de péage, chaque automobiliste doit payer 500 FCFA. Or cette tarification globale non seulement sous-estime la longueur du réseau utilisée par l'automobiliste mais de plus, elle n'est pas adaptée à la taille du véhicule. Certains axes routiers sont surtaxés arbitrairement par rapport à d'autres (Ex : Ngaoundéré - Garoua a 2 postes de péages pour 296 km, soit 1000 FCFA/296 km = 3,38 FCFA/km alors que Limbé -Idenau 11,36 FCFA/km (44 km et un poste de péage). La dégradation de la chaussée est fonction de l'agressivité qu'elle subit

de la part du trafic routier. Il est donc logique de taxer les usagers sur cette base. Les véhicules lourds détériorent beaucoup plus la chaussée que le font les véhicules légers. L'étude du système de taxation des véhicules routiers au Cameroun a révélé qu'un camion lourd détériore environ 2,8 fois plus la chaussée qu'une voiture particulière ou un taxi . D'où la nécessité de redéfinir la politique tarifaire du péage dans le sens d'une plus grande équité de manière à faire payer aux usagers les montants équivalents aux charges d'entretien et de fonctionnement qu'ils occasionneraient en utilisant le réseau.

- L'absence des itinéraires alternatifs : d'après le principe même de tout péage routier, l'existence des itinéraires alternatifs constitue l'une des conditions nécessaires à l'instauration d'un péage. Or par rapport aux axes bitumés à péage, il n'existe pas des itinéraires alternatifs, et donc par conséquent l'usager est obligé de payer.

1.3.2 Bilan fonctionnel

La gestion du péage au niveau central est du ressort de la Direction Générale des Impôts du Ministère des Finances. Elle centralise la distribution des tickets de péage. Du fait de l'absence d'un système informatique de gestion, il est difficile de connaître avec précision la demande en tickets dans les différents postes de péage. Bien plus, à cause de l'éloignement de certains postes par rapport à l'unité centrale de distribution, il se pose un problème de transport.

Au niveau local, la gestion du péage est assuré par les agents d'astreinte. Cette gestion, bien qu'en amélioration, reste médiocre, dû :

- au laxisme et aux malversations des agents de contrôle du péage (irresponsabilité et malhonnêteté, pratiques de réglementation particulières qui ignorent les textes en vigueur, détournement des recettes du péage) ;
- à l'incivisme de certains usagers qui, de connivence avec les agents de contrôle, se soumettent aux réglementations particulières afin de contourner le péage.

Pour les deux catégories de contrevenants cités ci-dessus, leurs comportements vis-à-vis du péage sont la conséquence de leur condition de travail. Celles-ci sont rudes, puisque les contrôleurs font face à plusieurs types de problèmes, notamment de transport, d'insécurité, de primes de risques et des indemnités promises à titre de motivation, d'équipements appropriés aux intempéries,...

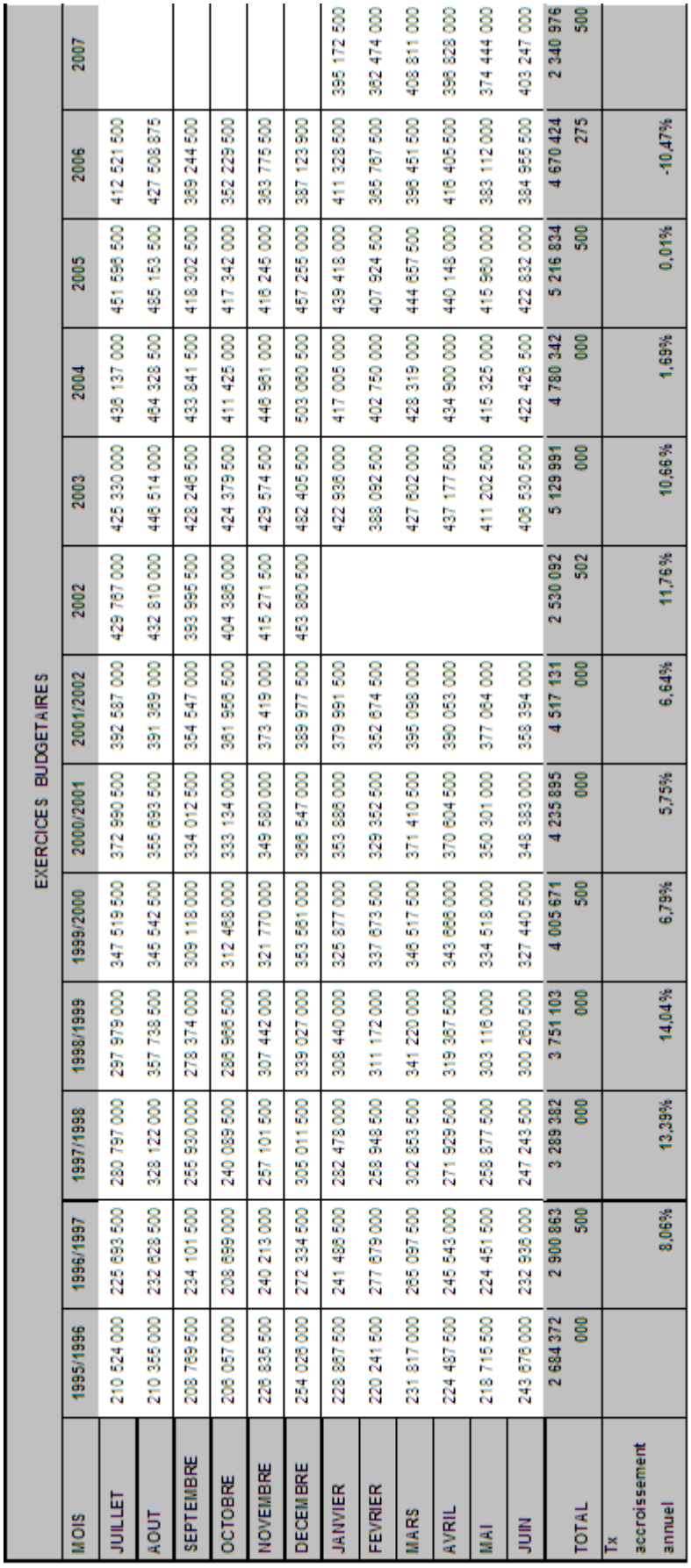
Le péage, dans sa formule actuelle, présente un grand nombre d'insuffisances. Sur le plan de l'organisation, on note une confusion entre les attributions du **CISOPR** et du **PSRR**. Cette confusion se fait également remarquer en ce qui concerne la concertation entre les différentes institutions impliquées dans le péage. Sur le plan du fonctionnement, deux insuffisances importantes retiennent l'attention. Il s'agit d'une part, de la mauvaise politique tarifaire et, d'autre part, de la gestion archaïque des postes et des fonds du péage. Cet ensemble de facteurs entrave considérablement l'efficacité de cette taxe. D'où l'urgence, à moyen ou long terme, de revoir en profondeur les aspects négatifs évoqués ci-dessus afin de mobiliser davantage notre péage routier, lui permettant ainsi de renouer avec l'efficacité et la rentabilité.

1.4 Problématique

Les projections des ressources du Fonds routier prennent en compte la cohérence des interventions des partenaires au développement. Elles se fondent sur les actions ci-dessous envisagées par le gouvernement :

- ◇ Le relèvement des ressources à affecter au Fonds routier, notamment par le relèvement de la RUR, et le versement effectif des recettes du péage, de la taxe à l'essieu et des amendes issues du péage ;
- ◇ L'amélioration du recouvrement des droits de péage et des amendes routières dans le cadre du Programme de sécurisation des Recettes Routières, chargé de toutes les opérations relatives au péage routier, notamment l'animation des postes de contrôle, la collecte et le suivi des recettes ;
- ◇ L'établissement en 2006 des modalités de relèvement à partir du 01^{er} janvier 2007 du tarif du péage routier en fonction de la capacité de dégradation de la route ;
- ◇ La possibilité d'élargir l'assiette des prélèvements liés à l'usage et à l'accès à la route au profit de Fonds Routier, notamment en ce qui concerne le transit routier.

L'entretien et la construction du réseau routier au Cameroun est évalué à environ 85 milliards de FCFA par an ; c'est ainsi que dans le cadre de l'atteinte du point d'achèvement de l'initiative PPTE, les partenaires au développement allouent annuellement à l'Etat camerounais la somme de 45 milliards de FCFA. La somme restante incombe à l'Etat et elle est répartie de telle sorte que le péage se devait de fournir 5,7 milliards de FCFA en 2006 avec une augmentation annuelle visant à atteindre 8 milliards de FCFA en 2012. Tel n'a pas été le cas jusqu'ici ; c'est la raison pour laquelle le péage routier camerounais se doit de mettre sur pied des stratégies visant à atteindre ces objectifs.



EXERCICES BUDGETAIRES														
MOIS	1995/1996	1996/1997	1997/1998	1998/1999	1999/2000	2000/2001	2001/2002	2002	2003	2004	2005	2006	2007
JUILLET	210 524 000	225 693 500	280 797 000	297 979 000	347 519 500	372 990 500	392 597 000	429 787 000	425 330 000	438 137 000	451 596 500	412 521 500	
AOUT	210 355 000	232 628 500	328 122 000	357 738 500	345 542 500	355 893 500	391 369 000	432 810 000	446 514 000	464 328 500	485 153 500	427 508 875	
SEPTEMBRE	208 769 500	234 101 500	255 930 000	278 374 000	309 118 000	334 012 500	354 547 000	393 995 500	428 246 500	433 841 500	418 302 500	389 244 500	
OCTOBRE	205 057 000	208 699 000	240 089 500	286 966 500	312 468 000	333 134 000	361 966 500	404 388 000	424 379 500	411 425 000	417 342 000	352 229 500	
NOVEMBRE	226 835 500	240 213 000	257 101 500	307 442 000	321 770 000	349 580 000	373 419 000	415 271 500	429 574 500	446 961 000	416 245 000	383 775 500	
DECEMBRE	254 026 000	272 334 500	305 011 500	339 027 000	353 561 000	388 547 000	389 977 500	453 880 500	482 405 500	503 060 500	457 255 000	397 123 500	
JANVIER	228 887 500	241 486 500	282 478 000	308 440 000	325 877 000	353 888 000	379 991 500		422 936 000	417 005 000	439 418 000	411 328 500	395 172 500
FEVRIER	220 241 500	277 679 000	258 948 500	311 172 000	337 673 500	329 352 500	352 674 500		388 092 500	402 750 000	407 924 500	385 787 500	382 474 000
MARS	231 817 000	265 097 500	302 853 500	341 220 000	346 517 500	371 410 500	395 088 000		427 602 000	428 319 000	444 657 500	396 451 500	408 811 000
AVRIL	224 487 500	245 543 000	271 929 500	319 367 500	343 669 000	370 604 500	390 053 000		437 177 500	434 900 000	440 148 000	416 405 500	395 828 000
MAI	218 715 500	224 451 500	258 877 500	303 116 000	334 518 000	350 301 000	377 064 000		411 202 500	415 325 000	415 960 000	383 112 000	374 444 000
JUIN	243 678 000	232 938 000	247 243 500	300 280 500	327 440 500	348 383 000	358 394 000		408 530 500	422 428 500	422 832 000	384 955 500	403 247 000
TOTAL	2 684 372 000	2 900 863 500	3 289 382 000	3 751 103 000	4 005 671 500	4 235 895 000	4 517 131 000	2 530 092 502	5 129 991 000	4 780 342 000	5 216 834 500	4 670 424 275	2 340 976 500
Tx accroissement annuel		8,06%	13,39%	14,04%	6,79%	5,75%	6,64%	11,76%	10,66%	1,69%	0,01%	-10,47%	

FIG. 1.2 – Recettes mensuelles du péage routier suivant les différents exercices budgétaires.

Chapitre 2

Présentation et d'escription des données

2.1 Présentation des données

Le but de notre étude est de trouver les moyens d'optimiser les recettes du péage routier. Pour le faire, nous tiendrons compte du Trafic Moyen Imposable (TMI) sur l'ensemble des postes de péage du Cameroun, du Volume du Parc Automobile Camerounais (VPAC), de l'évolution des : Prix du Gasoil à la Pompe (PGP), du Super à la Pompe (PSP), Prix Moyen du Carburant (PMC).

Les données concernant le Trafic Moyen Imposable (TMI) nous viennent du Ministère des Travaux Publics. Ils y disposent de données de comptage nécessaires à l'amélioration de la qualité des routes. Pour avoir le TMI sur l'ensemble des postes de péage nécessaire à notre étude, il nous a fallu situer sur la carte routière l'ensemble des postes de péage et d'en tirer les données de comptage sur les axes routiers concernés.

Le Volume du Parc Automobile Camerounais (VPAC) nous a été fourni par le Ministère des Transports qui gère le nombre de véhicules en circulation sur le territoire camerounais.

Quant aux Prix du Gasoil et du Super à la Pompe, ils ont été mis à notre disposition par la Caisse de Régulation des Prix des Hydrocarbures (CRPH).

Les recettes que nous avons sont collectées dans les différents postes de péage du pays de 1993 à Juin 2007. Pour les années budgétaires 1993/1994 et 1994/1995, elles sont annuelles et, à partir de 1995/1996, elles sont mensuelles. De plus, les années budgétaires de 1994/1995 à 2001/2002 vont de Juillet à Juin alors qu'à partir de 2003 elles vont de Janvier à Décembre avec une année transitoire de 6 mois de juin à décembre 2002.

Comme nous l'avons fait remarquer au chapitre précédent, vu la forte baisse des recettes de 2005 à 2006, nous nous arrêterons à l'année 2005 dans le cadre de notre étude et ainsi nos données iront de Juillet 1995 à Décembre 2005.

Les données sont stockées dans un tableau Excel de 114 lignes correspondant aux mois concernés par notre étude et 6 colonnes représentant les différentes variables : Recettes, TMI, VPAC, PSP, PGP, PMC. Toutes les autres variables de notre étude s'observent chaque mois avant les recettes : le trafic moyen imposable et les prix des carburants pour un mois quelconque sont connus au plus tard au début de ce mois, les données sur le volume du parc automobile quant à elles sont annuelles alors que les données sur les recettes sont disponibles au plus tôt le 1er du mois suivant.

Nous disposons également de recettes collectées dans chaque poste de péage pour les années 2003 à 2006 ; celles-ci nous seront utiles pour apprécier le poids de chacun des différents postes.

2.2 Description des données

Pour les calculs numériques, les graphiques, les prévisions et les simulations, nous utiliserons le logiciel R (version non commerciale de S-plus).

Toutes les commandes R utilisées sont disponibles en annexe. Nos données sont dans différents tableaux (disponibles en annexe) qui du fait que nous débutons notre étude en janvier 1995 présentent plusieurs données manquantes pour les variables recettes, TMI, PSP, PGP et donc PMC dans les proportions respectives : 0.02, 0.04, 0.04 et 0.04. Nous remplacerons ces données manquantes par les techniques d'imputation que nous présenterons au chapitre suivant. Ces données manquantes viennent du fait qu'avant l'exercice budgétaire 1995/1996 les recettes du péage ne sont pas notées mensuellement, aussi le trafic moyen sur les axes routiers du pays et l'évolution du prix du carburant ne sont pas ne sont pas archivés avant 1996.

La figure 2.1 suivante montre l'évolution des recettes mensuelles du péage routier camerounais de 1995 à 2007.

On remarque que jusqu'en 2005, la tendance de cette chronique est linéaire et varie très peu ; donc elle paraît essentiellement composée des variations saisonnières. Les trois figures qui suivent montrent les évolutions des recettes du péage routier de Juillet 1995 à Juin 2007, par année cette fois.

On tire de ces trois figures qu'au fil des années :

- Les mois de Décembre et Août présentent les recettes les plus élevées.
- Les mois de Février et d 'Octobre, quant à eux, sont ceux où les recettes sont au plus bas.

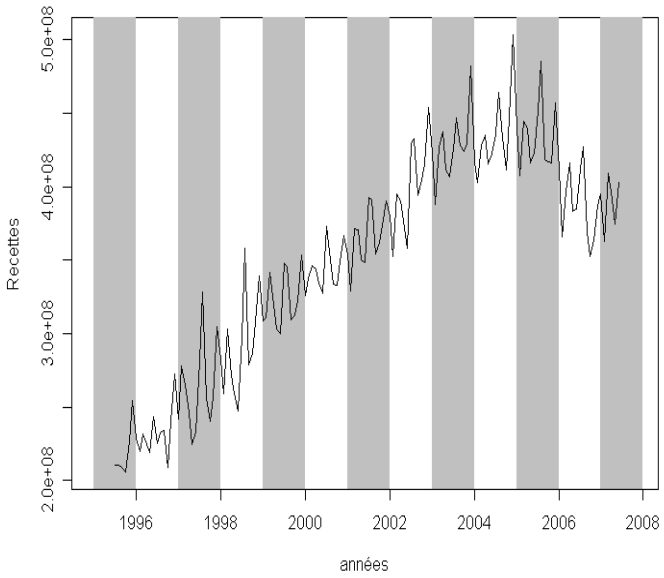


FIG. 2.1 – Evolution des recettes mensuelles du peage routier Camerounais.

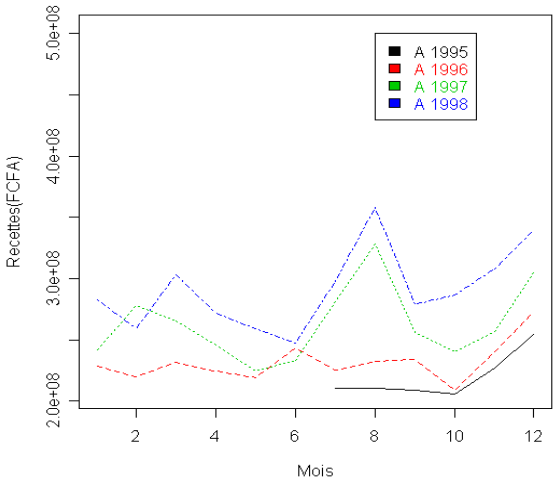


FIG. 2.2 – Evolution mensuelle des recettes du peage routier Camerounais de 1995 à 1998.

Les figures 2.2, 2.3 et 2.4 présentent les évolutions mensuelles des recettes des différents postes de péage du pays de 2003 à 2006.

Il ressort de l'observation des figures qui présentent l'évolution des recettes différents postes de péage du pays de 2003 à 2006 disponibles en annexe que :

- ⋆ les postes de péage les plus rentables du Cameroun sont ceux d'ÉDÉA, NKOMETOU, MBANGA, MBANKOMO, BOUMNYEBEL et TIKO. Ils rapportent à eux seuls près de 63% des recettes totales avec plus de 34% pour les postes d'ÉDÉA, NKOMETOU et MBANGA.
- ⋆ Les postes de péage qui rapportent le moins sont ceux de MAYO OULO, YAGOUA, NGATT et BEKASICHI. Ils ne rapportent que 0.7% de l'ensemble des

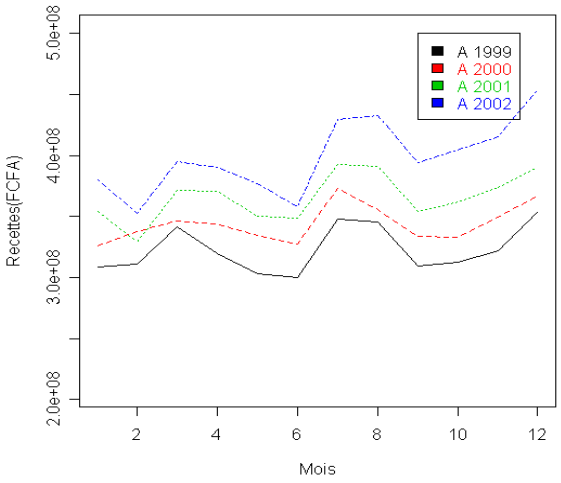


FIG. 2.3 – Evolution mensuelle des recettes du peage routier Camerounais de 1999 à 2002.

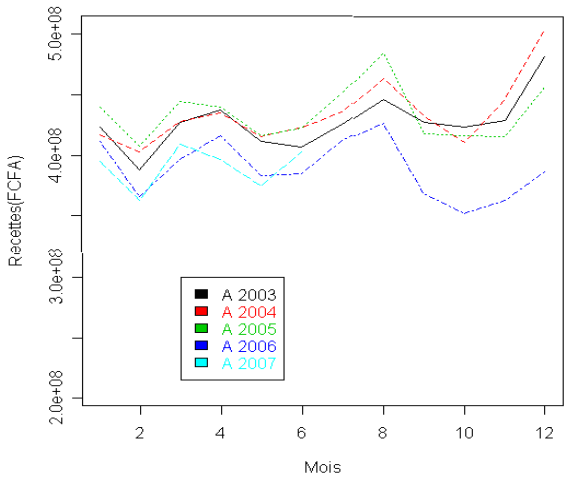


FIG. 2.4 – Evolution mensuelle des recettes du peage routier Camerounais de 2003 à 2007.

recettes.

Les autres données utiles à notre analyse sont également des séries temporelles dont les évolutions sont présentées dans graphiques 2.5, 2.6, 2.7 et 2.8.

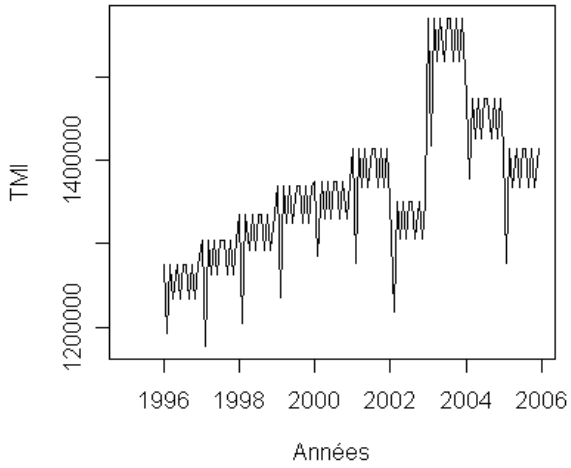


FIG. 2.5 – Evolution mensuelle du TMI.

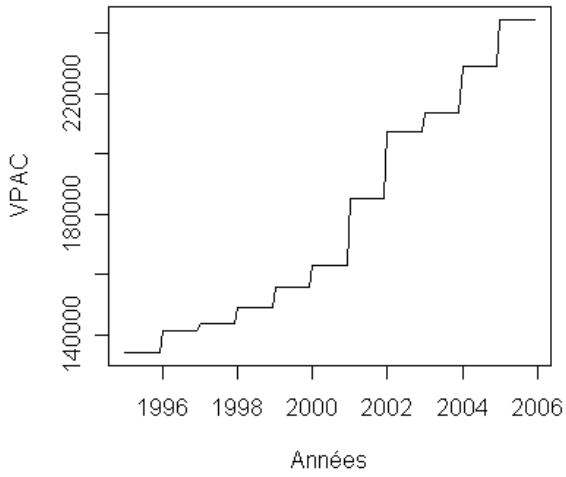


FIG. 2.6 – Evolution mensuelle du VPAC.

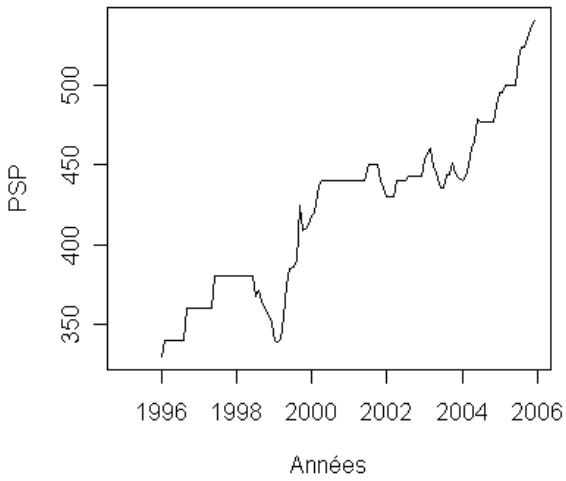


FIG. 2.7 – Evolution mensuelle du PSP.

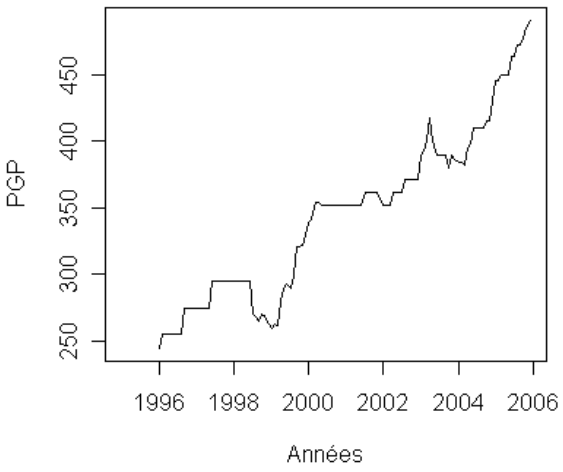


FIG. 2.8 – Evolution mensuelle du PGP.

Chapitre 3

Méthodes statistiques

Nous présentons dans ce chapitre les méthodes statistiques nécessaires à la réalisation de notre étude ; il s'agit des séries temporelles, car nos principales données qui sont les recettes collectées aux différents poste de péage du Cameroun varient avec le temps ; des techniques d'imputation , vu que certaines de nos données sont « manquantes » ; et de la régression linéaire multiple qui nous permettra de prédire les recettes du péage routier dans les années à venir en prenant en compte d'autres paramètres.

3.1 Séries chronologiques

Nous nous sommes inspirés ici de [3].

3.1.1 Modèles déterministes

Définition 3.1.1. *Une série chronologique* $\{Y_t\,,\, t \in T\}$ *est une suite d'observations d'une variable* Y *à différentes dates* t *indexées par un ensemble ordonné* T*. Habituellement,* T *est fini de sorte que* $T = \{t_1, t_2, \ldots, t_n\}$.

On supposera, dans toute la suite, que les dates sont équidistantes et donc, nous adopterons la notation simplifiée pour l'ensemble d'indices $T = \{1, 2, \ldots, n\}$. Ainsi, la série s'écrira $\{Y_t,\, t = 1, 2, \ldots, n\}$.

La représentation graphique des observations est une étape indispensable avant d'entreprendre une analyse plus technique de la chronique. Cette représentation permet d'apprécier l'évolution lente du phénomène (tendance), de dégager les périodes de stabilité. De ce qui précède se dégagent les notions de tendance, saisonnalité qui entrent dans la décomposition d'une série temporelle ou encore une chronique.

On considère qu'une série Y_t est la résultante de différentes composantes fondamentales :

- La **tendance** (ou trend) $\{C_t\}$ représente l'évolution à long terme de la série ;

- La **composante saisonnière** ou **saisonnalité** $\{C_t\}$ correspond à un phénomène qui se répète à intervalles de temps réguliers (phénomène périodique). En général, c'est un phénomène saisonnier, d'où le terme de variations saisonnières. La composante saisonnière est donc périodique de période p et il suffit de connaître ses p premières valeurs $S_1, , S_2, \ldots, \S_p$ (par périodicité, on a $S_t = S_t + p$ pour tout t).
- La **composante résiduelle** ou **bruit** ou **résidu** $\{\varepsilon_t\}$: représente les fluctuations irrégulières, en général de faible intensité, mais de nature aléatoire.

Les composantes ci-dessus peuvent se combiner selon différents modèles.

a- **Modèle additif**

$$Y_t = C_t + S_t + \varepsilon_t \quad \text{avec} \quad t = 1, 2, \ldots, n.$$

Hypothèses : $\sum_{j=1}^{p} S_j = 0$ et $\sum_{t=1}^{n} \varepsilon_t = 0$.

b- **Modèle multiplicatif**

$$Y_t = C_t \times S_t \times \varepsilon_t \quad \text{avec} \quad t = 1, 2, \ldots, n.$$

Hypothèses : $\sum_{j=1}^{p} S_j = p$ et $\frac{1}{n}\sum_{t=1}^{n} \varepsilon_t = 1$.

Pour effectuer l'analyse d'une série chronologique, on essaie d'abord de déterminer si les composantes de cette série peuvent être combinées selon un modèle additif ou un modèle multiplicatif. Afin de faire cette distinction, nous présentons une méthode d'identification décrite dans [3], en supposant que la série a une périodicité annuelle.

Méthode d'identification du type de schéma

1. On calcule d'abord les moyennes et les écarts-types pour chacune des périodes considérées. C'est-à-dire, par exemple pour une chronique ayant une périodicité annuelle, on calcule les moyennes et les écarts-types des observations pour chacune des années de cette chronique.
2. On calcule ensuite la droite des moindres carrés $\sigma = a\,\overline{x} + b$. C'est-à-dire, on calcule la droite de régression de ces écarts-types en fonction de ces moyennes.
3. Enfin, si a est significativement non différent de 0 à un seuil que l'on s'est fixé (par exemple 5%), nous pouvons conclure que le modèle de composition est additif, sinon, le modèle de composition est multiplicatif.

Définition 3.1.2. *Une série des moyennes mobiles d'ordre k, notée $MM(k)$, est la série des moyennes de k observations consécutives et elle prend ses valeurs aux dates moyennes correspondantes. Plus précisément, on calcule les moyennes de k termes consécutifs :*

1. *pour les dates :* $\frac{t_1+t_2+\ldots+t_k}{k}$, *puis* $\frac{t_2+t_3+\ldots+t_k+1}{k}, \cdots$, *jusqu'à* $\frac{t_{n-k}+t_{n-k-1}+\ldots+t_n}{k}$;
2. *et pour la variable d'intérêt :* $\frac{Y_1+t_2+\ldots+Y_k}{k}$, *puis* $\frac{Y_2+t_3+\ldots+Y_k+1}{k}, \cdots$, *jusqu'à* $\frac{Y_{n-k}+Y_{n-k-1}+\ldots+Y_n}{k}$.

Remarque 3.1.1. *Si k est impair : $k = 2\,m + 1$, la série moyenne mobile est calculée aux mêmes instants que les observations initiales. Ainsi, les dates de la série des moyennes mobiles sont des entiers naturels. En revanche, lorsque k est pair : $k = 2\,m$, la série moyenne mobile est calculée entre les dates des observations de la série initiale. Ainsi, les dates de la série des moyennes mobiles ne sont pas des entiers naturels.*

Ainsi, une moyenne mobile d'ordre pair se calcule à des dates qui ne coïncident pas avec les dates des observations. Si l'on veut comparer la série des moyennes mobiles avec la série initiale, on a besoin d'avoir des valeurs pour les mêmes dates d'observation. On définit les moyennes mobiles centrées pour pallier cet inconvénient des moyennes mobiles d'ordre pair.

Définition 3.1.3. *En gardant les notations ci-dessus, on définit la série des* ***moyennes mobiles centrées*** *d'ordre $k = 2\,m$ notée $MMC(k)$, par :*

$$MMC(k)_t = \frac{0.5\,Y_{t-m} + \ldots + Y_t + \ldots + 0.5\,Y_{t-m}}{2\,m}, \quad t = m+1, \ldots, n-m \ ;$$

et d'ordre $k = 2\,m + 1$ par :

$$MMC(k)_t = \frac{Y_{t-m} + \ldots + Y_t + \ldots + Y_{t-m}}{2\,m}, \quad t = m+1, \ldots, n-m,$$

où n est le nombre total des observations de la série initiale.

Remarque 3.1.2. *Si $k = 2\,m$ ou $k = 2, m+1$, m observations sont perdues à chaque extrémité de la série des moyennes mobiles centrées d'ordre k. Ainsi, la série initiale et la série des moyennes mobiles centrées d'ordre k n'ont pas la même longueur pour $k > 1$.*

Proposition 3.1.1. *Si une série chronologique à une composante saisonnière de période p, alors une moyenne mobile d'ordre p absorbe cette composante saisonnière, c'est-à-dire l'annule [7].*

Un algorithme de modélisation d'une chronique

Nous présentons dans cette partie un algorithme permettant de caractériser une chronique par un modèle déterministe. Pour simplifier l'exposé, nous présentons cette démarche en l'appliquant à une chronique dont la fréquence des observations est mensuelle.

Soit une série chronologique $\{Y_t\}_{t=1,\ldots,np} = \{Y_{ij}\}_{i=1,\ldots,n}^{j=1,\ldots,n}\}$, t est le nombre de mois à Partir de la date 0 ; i est le numéro de l'année ; j est le numéro du mois de l'année i.

1. On estime la tendance en éliminant la composante saisonnière à l'aide des moyennes mobiles centrées dont l'ordre est la période de la saisonnalité.

$$C_t = MMC(k)_t, \quad t = (i-1)+j, \quad i = 1, \ldots, n, \ j = 1, \ldots, p.$$

2. On identifie le modèle de composition : schéma additif ou multiplicatif.

   a- **Cas d'un modèle additif** :

   On exclut les valeurs extrêmes perdues lors du calcul de la série des moyennes mobiles centrées. Puis,

   – on calcule les données sans tendance (ou différences saisonnières) $Y_t - C_t$;

   – on calcule la moyenne des données sans tendance du mois j sur les n années, ceci pour chacun des mois. Ce sont les coefficients saisonniers. D'où

$$S_j = \frac{1}{n} \sum_{i=1}^{n} (Y_{ij} - C_{ij}) \ ;$$

   – On calcule la moyenne $\overline{S}$ des S_j :

$$\overline{S} = \frac{1}{p} \sum_{j=1}^{p} S_j \ ;$$

   Si $\overline{S} \neq 0$, on corrige les S_j : $S'_j = S_j - \overline{S}$.

   b- **Cas d'un modèle multiplicatif** :

   – On calcule les données sans tendance (ou rapports saisonniers) : $\frac{Y_t}{C_t}$;

   – on calcule la moyenne des données sans tendance du mois j sur les n années, ceci pour chacun des mois,

$$S_j = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_{ij}}{C_{ij}} \ ;$$

   – on calcule la moyenne des S_j : $\overline{S} = \frac{1}{p} \sum_{j=1}^{p} S_j$.
   Si $\overline{S} \neq 1$, on corrige les S_j : $S'_j = \frac{S_j}{\overline{S}}$

   On obtient à la fin de cette 2^{ime} étape, la série des variations saisonnières : pour tout i, $S_{ij} = S'_j$ ceci pour tous les mois j.

3. On calcule la série corrigée des variations saisonnières (CVS)

   a- **Cas d'un modèle additif** :

$$D_{ij} = Y_{ij} - S_{ij} = Y_{ij} - S'_j$$

   b- **Cas d'un modèle multiplicatif** :

$$D_{ij} = \frac{Y_{ij}}{S_{ij}} = \frac{Y_{ij}}{S'_j}$$

On peut réévaluer la tendance à l'issue de cette 3^{ime} étape par ajustement global de la série corrigée des variations saisonnières (CVS).

4. On calcule la série ajustée
   a- **Cas d'un modèle additif** :

$$\widehat{Y}_t = C_t + S_t \quad \text{ou} \quad \widehat{Y}_t = C_{ij} + S'_j.$$

   b- **Cas d'un modèle multiplicatif** :

$$\widehat{Y}_t = C_t \times S_t \quad \text{ou} \quad \widehat{Y}_t = C_{ij} \times S'_j.$$

5. Calcul des variations accidentelles ou résiduelles
   a- **Cas d'un modèle additif** :

$$\varepsilon_t = Y_t - \widehat{Y}_t.$$

   b- **Cas d'un modèle multiplicatif** :

$$\varepsilon_t = \frac{Y_t}{\widehat{Y}_t}.$$

On peut affiner les choses en recherchant la structure du bruit ε_t suivant les techniques des pages suivantes

a- **Pour un modèle additif**, on a : $Y_t = C_t + S_t + \varepsilon_t$.

b- **Pour un modèle multiplicatif**, on a : $Y_t = C_t \times S_t \times \varepsilon_t$.

Le modèle étant retenu, on peut faire des prévisions très facilement. On prévoit la tendance en calculant C_{np+1} , C_{np+2} , ..., puis, selon le modèle de composition, on ajoute ou on multiplie par le coefficient saisonnier du mois correspondant.

3.1.2 Modèles stochastiques

Soit (Ω, A, P), un espace probabilisé. (T, Γ) et (Ω', A') deux espaces mesurables.

Définition 3.1.4. *Un processus stochastique est une application X définie sur $\Omega \times T$, à valeurs dans Ω' associant au couple (ω, t) la réalisation $X(\omega, t)$, encore notée $X_t(\omega)$, et tel que pour t fixé appartenant à T, X est une variable aléatoire (v.a.) sur (Ω, A). Par extension, on écrira un processus sous la forme d'une famille de v.a. indicées par t notée $(X_t, t \in T)$ ou, plus simplement (X_t). La loi du processus est l'image P^X de P par X.*

Lorsque $\Omega' = \mathbb{R}$, le processus est dit unidimensionnel ou univarié.
Lorsque $T = \mathbb{Z},$, le processus est dit en temps discret.

Nous considérerons dans la suite les processus stochastiques univariés et à temps discret.

Définition 3.1.5. *Les **modèles de prévision** sont les modèles qui cherchent à un instant donné t, à « prévoir » pour les instants $t + \Delta t$ le devenir d'une réalisation connue jusqu'à t.*

Définition 3.1.6. *Les **modèles de simulation** sont des modèles qui ne cherchent pas à reproduire une partie de la réalisation d'un processus, mais à générer des scénarios « possibles » d'un processus dont on connaît une réalisation. Leur principe est d'utiliser les générateurs de variables aléatoires qui devront respecter la structure statistique des processus à reproduire.*

Définition 3.1.7. *L'**opérateur retard** B est un opérateur qui, à un processus X_t, associe le processus Y_t tel que $Y_t = B\,X_t = X_{t-1}$.*

Définition 3.1.8. *On dit que X_t est strictement (ou fortement) stationnaire si pour toute suite finie d'instants $t_1, t_2, \ldots, t_k$ éléments de $\mathbb{Z}$ et tout entier $r \in \mathbb{Z}$ les lois jointes de $(X_{r+t_1}, X_{r+t_2}, \ldots, X_{r+t_k})$ et de $(X_{t_1}, X_{t_2}, \ldots, X_{t_k})$ sont les mêmes (lois jointes invariantes par translation dans le temps).*

Définition 3.1.9. *Un processus X_t est stationnaire au second ordre (ou faiblement stationnaire) si ses moyennes et ses covariances sont invariantes par translation. C'est-à-dire :*

$$\begin{cases} \mu_t = E[X_t] = \mu \\ \gamma(t,j) = E[(X_{t+j} - \mu)(X_t - \mu)] = \gamma(j) \quad \textit{pour tout} \quad t, j \in \mathbb{Z} \end{cases}$$

Dans la suite, stationnaire signifiera stationnaire au second ordre. Supposons que X_t est un processus stationnaire au second ordre.

Définition 3.1.10. *γ est appelé fonction d'autocovariance.*

Remarque 3.1.3. *$Var(X_t) = \gamma(0)$, pour tout $t \in \mathbb{Z}$.*

Théorème 3.1.1. *(i)* $\gamma(j) = \gamma(-j), \ \forall\, j \in \mathbb{Z}$

(ii) $\gamma(0) \geq 0$.

(iii) γ est une fonction définie positive, c'est-à-dire :

$$\sum_{j=1}^{n}\sum_{k=1}^{n} (\gamma(t_k) - t_j)\, z_j\, z_k \geq 0\,, \forall\, n \geq 0 \forall\, (t_i)_{i=1}^{n}\, (z_j)_{j=1}^{n} \in \mathbb{R}^n$$

On pose pour tout $j \in \mathbb{Z}$, $\rho(j) = \frac{\gamma(j)}{\gamma(0)}$.

Définition 3.1.11. *ρ est appelé fonction d'autocorrélation.*

Corollaire 3.1.1. *La fonction d'autocorrélation ρ a toutes les propriétés de la fonction d'autocovariance et satisfait la condition supplémentaire $\rho(0) = 1$.*

On pose :

$$r(k) = \frac{Cov\left(X_1 - P_{[X_2,X_3,\ldots,X_k]}(X_1)\,, X_{k+1} - P_{[X_2,X_3,\ldots,X_k]}(X_{k+1})\right)}{Var\left(X_1 - P_{[X_2,X_3,\ldots,X_k]}(X_1)\right)}, \quad \text{pour} \quad k \in \mathbb{Z},$$

où $P_{[X_2,X_3,\ldots,X_k]}(X_1)$ et $P_{[X_2,X_3,\ldots,X_k]}(X_{k+1})$, sont respectivement la régression ou la projection de X_1 (respectivement de X_{k+1}) sur le sous-espace engendré par les variables aléatoires X_1, $X_2, \ldots, X_k$.

Définition 3.1.12. *r est appelé fonction d'autocorrélation partielle.*

Définition 3.1.13. *Un bruit blanc est un processus ε_t, $t \in \mathbb{Z}$ centré tel que :*

$$E[\varepsilon_t\,\varepsilon_s] = \sigma^2\,\delta_s^t \quad avec \quad \sigma > 0 \ et \ \delta_s^t = \begin{cases} 1, & si\ t = s \\ 0, & sinon \end{cases}$$

Proposition 3.1.2. *Un estimateur empirique de la fonction d'autocorrélation (ACF) ρ est définie par : $\widehat{\rho}_k = \frac{\widehat{\gamma(k)}}{\widehat{\gamma(0)}}$,*

$$o\grave{u} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_t \quad et \quad \widehat{\gamma(k)} = \begin{cases} \frac{1}{n-k}\sum_{i=1}^{n-k}(X_{t+k}-)\overline{X})(X_t - \overline{X}), & si\ 0 \le k \le n-1 \\ 0, & sinon \end{cases}$$

Proposition 3.1.3. *Un estimateur empirique de la fonction d'autocorrélation partielle (PACF) s'obtient en résolvant le système de Yule-Walker suivant, en $(\alpha_1, \alpha_2, \ldots, \alpha_k)$,*

$$\widehat{\rho}(j) = \alpha_1\,\widehat{\rho}(j-1) + \alpha_2\,\widehat{\rho}(j-2) + \cdots + \alpha_k\,\widehat{\rho}(j-k), \quad j = 1, 2, \ldots, k$$

et en prenant $\widehat{r} = \alpha_k$, $k \in \mathbb{Z}$.

Définition 3.1.14. *X_t est un processus $ARMA(p,q)$ (autorégressif moyenne mobile) s'il est stationnaire et vérifie :*

$$\phi(B) = 1 - \phi_1\,B - \phi_2\,B^2 - \ldots - \phi_p\,B^p \neq 0 \quad \theta(B) = 1 + \theta_1\,B + \theta_2\,B^2 + \ldots + \theta_p\,B^p \neq 0,$$

où ϕ et θ sont des polynômes de degré p et q respectivement. et dont les racines sont de module supérieur à 1 et ne sont pas communes.

Définition 3.1.15. *Un processus auto régressif d'ordre p , noté $AR(p)$, est un processus $ARMA(p,q)$ avec $q = 0$.*

Définition 3.1.16. *Un processus moyenne mobile d'ordre q , $MA(q)$, est un processus $ARMA(p,q)$ avec $p = 0$.*

Définition 3.1.17. *Un processus X_t est intégré d'ordre d si les processus $(1-B)^n\,X_t\,, n = 1, 2, \ldots, d-1$, ne sont pas asymptotiquement équivalents à un processus stationnaire, mais la série $Y_t = (1-B)^d\,X_t$ l'est.*

Définition 3.1.18. *X_t est un processus $ARIMA(p,d,q)$ (autorégressif moyenne mobile intégré) s'il vérifie une équation du type : $\phi(B)\,\Delta^d X_t = \theta(B)\,\varepsilon_t$ $t \in t \in \mathbb{N}$ où ε_t est un bruit blanc, $\Delta^d X_t = (1-B)^d X_t$,*

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \neq 0 \quad \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_p B^p \neq 0,$$

θ e ϕ sont des polynômes de degré q et p respectivement et donc les racines sont de module supérieur à 1 et où les conditions initiales

$$Z_{-1} = \{X_{-1}, \ldots, X_{-p}\,\varepsilon_{-1,\ldots\,\varepsilon_{-q}}\},$$

sont non corrélées avec $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_t, \ldots$

Proposition 3.1.4. *Si X_t est un processus $ARIMA(p,d,q)$, alors le processus $(1-B)^d X_t$ est asymptotiquement un $ARMA(p,q)$.*

Remarque 3.1.4. *Les processus définis ci-dessus ont la forme générale suivante : $\Phi^* X_t = \Theta^*(B)\,\varepsilon_t$, où Θ^* et Φ^* sont des polynômes et ε_t est un bruit blanc.*

Définition 3.1.19. *Φ^* est appelé polynôme autorégressif du processus.*

Proposition 3.1.5. *Si X_t est un processus $AR(p)$, alors $r(k) = 0$ si $k > p$ et $r(p) \neq 0$.*

Proposition 3.1.6. *Si X_t est un processus $MA(q)$, alors $\rho(h) = 0$ si et $\mid h \mid > q$ et $\rho(q) \neq 0$.*

Méthodologie : modèle $ARIMA(p,d,q)$

On dispose des observations $x_1, \ldots, x_T$ de $X_1, \ldots, X_T$. Comment modéliser par un modèle $ARIMA(p,d,q)$?

Identification à priori et estimation

Première phase de l'identification : choix de d

Approche empirique : l'autocorrélogramme

Si les $\widehat{\rho}_T$ restent proches de 1 ou décroissent lentement avec h, alors le processus est sans doute non stationnaire.

Remarque 3.1.5. *Si l'autocorrélogramme fait penser que X_t est non stationnaire, alors on étudie l'autocorrélogramme du processus $Y_t = (1-B)\,X_t$, etc...*

Approche par test de racine unité : test de Phillips-Perron

L'hypothèse nulle est l'hypothèse de non stationnarité dans la série étudiée. La non stationnarité ici est caractérisée par la présence d'une racine unité dans le polynôme autorégressif du modèle.

Remarque 3.1.6. *Si ce test permet de ne pas rejeter l'hypothèse nulle, alors X_t est intégré d'ordre au moins 1. On effectue à nouveau ce test avec le processus $Y_t = (1 - B)\, X_t$, etc...*

La valeur de d est celle pour laquelle le processus $Y_t = (1-B)^d\, X_t$ est stationnaire.

Deuxième phase de l'identification : choix de p et q

On suppose que l'on a déjà d et on travaille éventuellement sur $Y_t = (1 - B)^d\, X_t$. On assimile a un processus $ARMA(p, q)$. On se propose donc de déterminer p et q.

Approche empirique : on cherche ici à déterminer la valeur de P et Q telle que Y_t est un $ARMA(P, 0)$ et un $ARMA(0, Q)$, ce qui est à peu près équivalent pour K grand à $r(k) = 0$ et $\rho(h) = 0$ pour $P + 1 \leq k \leq K$et $Q + 1 \leq h \leq K$. (K est le nombre maximal des autocorrélations que l'on désire calculer). Une fois P et Q déterminés, Box et Jenkins proposent en général de traiter séparément toutes les possibilités des couples (p, q) majorés par (P, Q).

Estimation

A l'issue des phases précédentes, on a choisi d et divers couples (p, q) compatibles avec les données. Le modèle s'écrit :

$$\phi(B)\, \Delta^d\, X_t = \theta(B)\, \varepsilon_t,$$

où ε_t est un bruit blanc de variance σ^2 et $\theta(B) = 1 + \theta_1\, B + \theta_2\, B^2 + \ldots + \theta_p\, B^p$, $\phi(B) = 1 - \phi_1\, B - \phi_2\, B^2 - \ldots - \phi_p\, B^p$.

Les paramètres à estimer sont : ϕ_1, ϕ_2, $\ldots$, ϕ_p, θ_1, θ_2, $\ldots$, θ_q. Pour obtenir des informations sur les méthodes d'estimation de ces paramètres (cf [1]).

3.1.3 Vérification à posteriori et choix du modèle

Vérification à posteriori

- **Tests sur les paramètres** : Dans cette partie, on effectue généralement le test $H_0 : \phi_p = 0$ contre l'hypothèse alternative $H_1 : \phi_p \neq 0$, et le test $H_0 : \theta_q = 0$ contre l'hypothèse alternative $H_1 : \theta_q \neq 0$. Si $\widehat{\phi}_p$ (ou $\widehat{\theta}_q$) n'est pas significatif, on relance l'estimation en remplaçant p par $p - 1$ (ou q par $q - 1$). Pour plus de précisions sur ces tests (cf [1]).
- **Test sur les résidus** : Les résidus estimés (à savoir $\widehat{\varepsilon}_t$) sont-ils compatibles avec l'hypothèse de bruit blanc de ε_t ? pour cela, on effectue le test du Portmanteau proposé par Box-Pierce dont la statistique a été améliorée par Ljung-Box. Ce test s'écrit :

H_0 : $\widehat{\varepsilon}_t$: est un bruit blanc contre l'hypothèse alternative : H_1 : $\widehat{\varepsilon}_t$ n'est pas un bruit blanc.

Les détails théoriques de ce test se trouvent dans [1].

Choix du modèle

A l'issue des phases d'estimation et de vérification, il reste en général plusieurs modèles possibles pour représenter les données. Nous choisirons dans cette étude le modèle pour lequel la variance des résidus σ^2 est la plus petite. On montre que l'erreur de prévision diminue avec cette variance [3]. Nous fondons notre choix sur ce critère pour la simple raison que l'objectif principal visé dans cette étude est de déterminer un modèle permettant de faire de « bonnes » prévisions.

3.2 Techniques d'imputation

Corriger la non réponse dans une enquête ou un essai clinique n'est jamais chose simple et évidente. Des techniques de correction existent, en particulier la répondération et l'imputation. Cependant pour être appliquées correctement, elles nécessitent de tenir compte du contexte dans lequel on se trouve. En effet, leur utilisation de façon mécanique permettrait effectivement de se sortir de n'importe quelle situation, mais avec le risque d'introduire du biais. L'avantage de la technique d'imputation est qu'elle permet d'obtenir des bases de données complètes, ce qui a l'avantage de préserver toute l'information sur les données et d'effectuer des analyses en utilisant des logiciels qui nécessitent les données complètes. Faisant référence à [5], nous noterons qu'on distingue 2 types d'imputation : l'imputation simple et l'imputation multiple.

- **Imputation simple**.
  C'est une technique d'imputation qui permet de remplacer une valeur manquante par une valeur plausible prédite ou simulée ; mais cette technique présente un réel inconvénient car elle ne reflète pas toute l'incertitude des valeurs manquantes.
- **Imputation multiple**.
  Création de plusieurs valeurs plausibles d'une donnée manquante. Le but n'est pas de :
- prédire avec la plus grande précision les données manquantes,
- décrire les données de la meilleure façon possible, Les buts sont :
- décrire correctement l'incertitude due aux données manquantes,
- préserver les aspects importants des distributions,
- préserver les relations importantes entre les variables.

Nous nous intéresserons ici aux différents types d'imputation simple vu que le taux de nos données manquantes est faible.

3.2.1 Imputation par la moyenne

Considérons la variable aléatoire $(X_1, \ldots, X_q, X_{q+1}, \ldots, X_n)$ où $(X_1, \ldots, X_q)$ représente les valeurs totalement observées et $(X_q, \ldots, X_n)$ les valeurs manquantes ; alors chaque valeur manquante est remplacée par :

$$\overline{X}_{obs} = \frac{1}{q} \sum_{i=1}^{q} X_i.$$

3.2.2 Utilisation d'un modèle de régression

Remplacement de chaque valeur manquante par une valeur prédite $(X_1, \ldots, X_n)$ valeurs observées $(Y_1, \ldots, Y_q)$ valeurs observées et $(Y_{q+1}, \ldots, Y_n)$ valeurs manquantes. On affectera alors aux valeurs manquantes des valeurs prédites par le modèle de régression linéaire Y X. Dans ce cas, les corrélations sont augmentées. Remplacement de chaque valeur manquante par une valeur prédite par le modèle plus un résidu aléatoire. Imputer où $\widehat{Y}_i + e_i$ où e_i $N(0, S^2)$, avec S^2 représentant l'erreur moyenne quadratique.

3.2.3 Imputation par une valeur observée tirée au hasard

- Préserve la distribution marginale de la variable.
- Peut fausser les corrélations avec d'autres variables.
- Appropriée pour des analyses unidimensionnelles.

3.3 Régression linéaire multiple

L'objet même de la régression est précisément l'étude, à partir d'un échantillon d'observations aléatoires, de la liaison stochastique entre la variable y dépendante et une variable x indépendante et certaine dans la population d'où a été tirée l'échantillon.

La démarche décrite dans [6] correspondante revient essentiellement à :

- Etudier à l'aide des informations d'un échantillon, la régression, en spécifiant le type de régression de y en x c'est-à-dire le lieu géométrique de la moyenne conditionnelle $\overline{y}_x$ en fonction de x, et en précisant la variabilité de y autour de cette courbe ; utiliser cette analyse de la régression en vue de pouvoir répondre convenablement à certaines questions de signification, comparaison, prédiction, qu'on peut être amené à se poser relativement à la population d'où est tirée l'échantillon.
- La régression multiple généralise la régression simple en étudiant la liaison stochastique entre une variable aléatoire y (la variable dépendante et p variables

Numéro d'observation	y	x_1	x_2	$\cdots$	x_p
1	y_1	x_{11}	x_{21}		x_{p1}
2	y_2	x_{12}	x_{22}		x_{p3}
3	y_3	x_{13}	x_{23}		
$\vdots$					
n	y_n	x_{1n}	x_{2n}	$\cdots$	x_{pn}

certaines $x_1, \ldots, x_p$ (les variables indépendantes) au sein d'une population donnée dont on observe un échantillon aléatoire. On suppose en outre que toutes les variables indépendantes x_j sont mesurées sans erreur.

Description des données et modèle

Au lieu d'avoir un régresseur comme c'est le cas de la régression linéaire simple, on en a p, qui sont notés $x_1, \ldots, x_p$. Les données se présentent sous la forme de n ensembles d'observations de la variable y et des p régresseurs :

Modèle : $y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + \mu_i \qquad i = 1, 2, \cdots$

Ajustement du modèle

Comme pour la régression linéaire simple, un des usages de la régression linéaire multiple consiste à prédire la valeur d'un y Comme pour la régression linéaire simple, un des usages de la régression linéaire multiple consiste à prédire la valeur d'un pour un ensemble de valeurs $x_1, \ldots, x_p$ données. La mesure de l'ajustement du modèle aux données est donc importante.

R^2 ne prend pas en compte le nombre de variables explicatives. C'est pourquoi on s'intéresse plutôt au $R^2 - Adj$, qui représente une mesure de l'ajustement corrigé par le nombre de régresseurs du modèle.

Sélection de variables explicatives

La sélection de variables est une perspective naturelle à plus d'un titre. La principale raison est que soit certaines variables ne contribuent pas à l'explication de la variable à expliquer, soit des variables sont très corrélées et apportent donc une redondance d'information. Dans ces deux situations, on a envie de les éliminer du modèle. Il faut bien noter qu'on recherche toujours à privilégier le modèle le plus simple possible permettant ainsi une interprétation facile. De plus, un trop grand nombre de variables peut mener à une augmentation de la variance résiduelle puisque le nombre de degrés de liberté lui, diminue.

L'objectif est donc de déterminer à partir de toutes les variables explicatives, un sous-ensemble de variables suffisamment explicatif.

Une première possibilité « brutale » consiste à évaluer toutes les régressions possibles. Malheureusement, cette solution est souvent très longue, voire impossible temporellement, dès lors que le nombre de variables est grand (le nombre de régressions étant 2^p). Différentes autres méthodes de sélection sont utilisées :

a- **Méthode ascendante** : la procédure commence avec le terme constant β_0, soit le modèle nul : $Y_i = \beta_0 + E_i$. Ensuite, elle s'effectue en plusieurs étapes :

Etape 1 : On choisit la variable x_{k1} parmi l'ensemble des variables de départ, qui contribue le plus à expliquer Y, i.e. celle qui fait augmenter le R^2 ou encore telle que $\widehat{\rho}(Y, X_{k1})$ est maximal. Ensuite, on teste la nullité du coefficient de régression associé et la variable est retenue en cas de significativité du test.

Etape 2 : On choisit la variable x_{k2} parmi l'ensemble des variables auquel on a retiré x_{k1}, telle que $\widehat{\rho}(Y,\ X_{k2}/X_{k1})$ est maximal. Ce n'est donc pas la variable la plus corrélée à Y, mais c'est celle qui apporte le plus d'informations en plus de x_{k1}. De la même façon que précédemment, le coefficient de régression est testé. Il existe plusieurs tests d'arrêt de la procédure : en choisissant un nombre à priori de variables ou une valeur finale de R^2, ou encore dès que le test de nullité de la dernière variable introduite n'est pas significatif.

b- **Méthode descendante** : C'est la procédure symétrique de la précédente, qui part du modèle complet et élimine à chaque étape la variable correspondant au plus petit coefficient de corrélation partielle.

c- **Méthode progressive** : Cette procédure est semblable à l'ascendance avec remise en cause à chaque étape des variables déjà introduites. En effet, il arrive souvent que des variables introduites en tête, par le biais de leur liaison avec une ou plusieurs autres variables introduites ultérieurement, ne soient plus significatives.

Chapitre 4

Applications et résultats.

4.1 Prévisions temporelles

Pour modéliser les recettes mensuelles du péage routier camerounais, la figure 2.1 du chapitre 2 nous suggère d'utiliser un schéma de composition additif. Nous le ferons en partant d'une désaisonnalisation de cette chronique à l'aide de la méthode des moyennes mobiles centrées tel que décrit dans la section I.2 du chapitre 3. La droite de régression des écarts-types en fonction des moyennes des recettes de chacune des années étudiées a pour équation

$$\sigma = 1.51 \times 10^{-2}\,\overline{x} + 1.8 \times 10^{7},$$

où le coefficient 1.51×10^{-2} n'est pas significatif (sa probabilité critique est 0.5655), (voir commande C 37 pour l'obtention de ces résultats). Ce qui suggère d'adopter effectivement un schéma de composition de type additif pour ces recettes mensuelles. La forme générale de ce type de modèle est :

$$Y_t = C_t + S_t + \varepsilon_t,$$

Où ε est la série des recettes que nous étudions, C_t est la composante tendancielle, S_t est la composante saisonnière, ε_t est la composante résiduelle.

4.1.1 Prévision simple : Tendance générale

Comme nous pouvons le voir par la Figure 4.1, de 1995 à 2005 les recettes suivent une tendance essentiellement linéaire donnée par l'équation de la droite (obtenue par la méthode des moindres carrés) :

$$\widehat{Y_t} = 2023667\,t + 204417147, \tag{4.1}$$

où t représente le numéro du mois, le mois n°1 étant Juillet 1995. (Voir commande C 38)

Les prévisions que nous pouvons faire par cette méthode pour les années de 2006 à 2008 sont consignées dans le tableau 4.1 :

Mois	Prévisions	2,5%	97,50%
janv-06	473564866	426012839	521116893
févr-06	475588533	428018820	523158246
mars-06	477612200	430024532	525199869
avr-06	479635867	432029974	527241761
mai-06	481659534	434035147	529283922
juin-06	483683202	436040051	531326352
juil-06	485706869	438044687	533369050
août-06	487730536	440049055	535412016
sept-06	489754203	442053155	537455250
oct-06	491777870	444056988	539498752
nov-06	493801537	446060554	541542520
déc-06	495825204	448063853	543586555
janv-07	497848871	450066885	545630857
févr-07	499872538	452069652	547675424
mars-07	501896205	454072153	549720257
avr-07	503919872	456074388	551765356
mai-07	505943539	458076359	553810719
juin-07	507967206	460078065	555856348
juil-07	509990873	462079506	557902240
août-07	512014540	464080684	559948396
sept-07	514038207	466081599	561994816
oct-07	516061874	468082250	564041499
nov-07	518085542	470082638	566088445
déc-07	520109209	472082764	568135653
janv-08	522132876	474082628	570183123
févr-08	524156543	476082230	572230855
mars-08	526180210	478081572	574278848
avr-08	528203877	480080652	576327102
mai-08	530227544	482079472	573375616
juin-08	532551211	484078032	580424390
août-08	534274878	486076332	582473424
août-08	536298545	488074373	584522717
sept-08	538322212	490072155	586572269
oct-08	540345879	492069679	588622079
nov-08	542369546	494066945	590672147
déc-08	544393213	496063954	592722473

TAB. 4.1 – Prévisions par (4.1) des recettes du péage de 2006 à 2008.

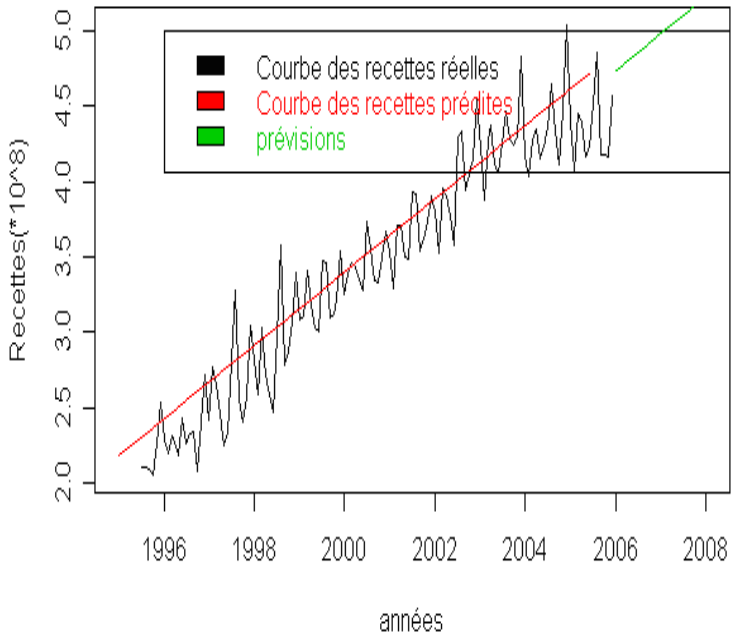


FIG. 4.1 – Tendance des recettes du peage routier Camerounais.

4.1.2 Prévision corrigée : en tenant compte des variations saisonnières

Dans notre étude, nous identifié un mouvement saisonnier assez clair dans les recettes du péage routier camerounais ; en tenant compte de ce mouvement saisonnier, on peut faire des prévisions à priori plus précises de l'évolution de ces recettes dans le temps en s'appuyant cette fois ci sur le modèle (4.2) suivant qui corrige (4.1) :

$$\widehat{Y}_t = 20233667\,t + 204417147 + S_t, \tag{4.2}$$

où S_t représente le coefficient saisonnier du mois t dans l'année calendaire. Ces coefficients saisonniers calculés en utilisant les moyennes mobiles sont donnés dans le tableau 4.2 :

Ainsi, afin d'expurger cette série de ses variations saisonnières, nous avons utilisé comme technique mathématique la méthode des moyennes mobiles qui a l'avantage de ne faire aucune hypothèse à priori sur la forme de la tendance à estimer. L'ordre de la moyenne mobile que nous utilisons pour désaisonnaliser cette chronique est de 12 correspondant à la période cette série qui est de 12 mois. La figure 4.3 représente la courbe décrite par la série des moyennes mobiles centrées d'ordre 12 et celle de la série des débits naturels. Nous obtenons cette figure par la commande C 39.

La figure 4.3 (obtenue par la commande C.40) présente la série de nos prévision tenant compte de la tendance et de l'estimation des coefficients saisooniers, ainsi qu'une

Mois	S_t
Janvier	13359120
Février	27731668.7
Mars	-11075940.4
Avril	-17355574.4
Mai	-1730671.6
Juin	32577011.2
Juillet	-1353520
Août	-14890682
Septembre	9899669.6
Octobre	474661.2
Novembre	-18029449.2
Décembre	-19606292.9

TAB. 4.2 – Coefficients saisonniers des différents mois.

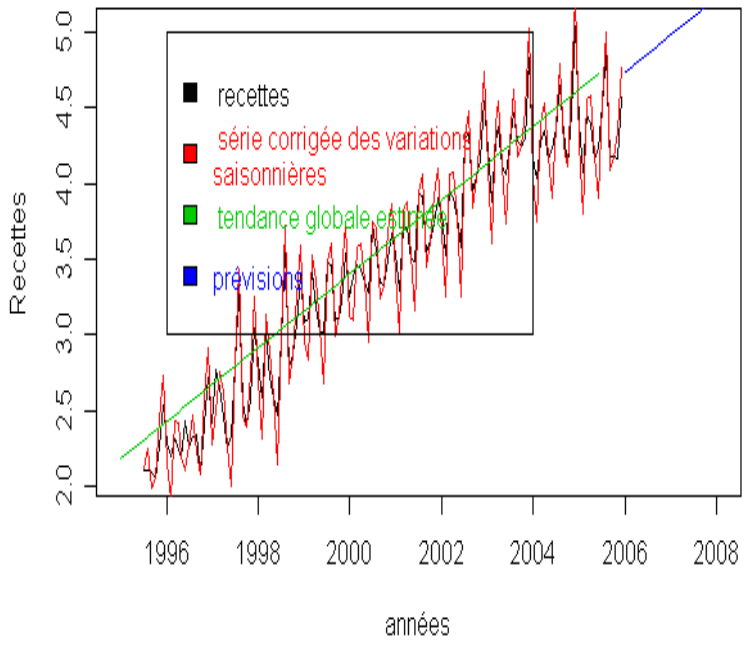


FIG. 4.2 – Série des moyennes mobiles d'ordre 12.

estimation de la tendance globale à long terme de la chronique des recettes.

Des prévisions plus précises pour les recettes du péage routier de janvier 2006 à Décembre 2008 sont fournies dans le tableau 4.3.

Mois	Prévisions	2,50%	97,50%
janv-06	474138135	404112926	544163345
févr-06	476165393	406114138	546216647
mars-06	478192650	408114954	548270346
avr-06	480219908	410115373	550324442
mai-06	4822447165	412115396	552378933
juin-06	484274422	414115024	554433821
juil-06	486301680	416114256	556489103
août-06	488328937	418113093	558544781
sept-06	490356195	420111537	560600852
oct-06	492383452	422109586	562657318
nov-06	494410709	424107242	564714176
déc-06	496437967	426104505	566771428
janv-07	498465224	428101377	568829072
févr-07	500492481	430097856	570887107
mars-07	502519739	432093944	572945534
avr-07	504546996	434089641	575004351
mai-07	506574254	436084948	577063559
juin-07	508601511	438079866	579123156
juil-07	510628768	440074394	581183143
août-07	512656026	442068534	583243518
sept-07	514683283	444062285	585304281
oct-07	516710541	446055649	587365432
nov-07	518737798	448048626	589426970
déc-07	520765055	450041217	591488894
janv-08	522792313	452033422	593551204
févr-08	524819570	454025242	595613899
mars-08	526846828	456016677	597676978
avr-08	528874085	458007728	599740442
mai-08	530901342	459998395	601804290
juin-08	532928600	461988680	603868520
août-08	534955857	463978582	605933132
août-08	536983115	465968103	607998127
sept-08	539010372	467957242	610063502
oct-08	541037629	469946001	612129258
nov-08	543034887	471934380	614195394
déc-08	545092144	473922380	616261908

TAB. 4.3 – Prévisions par (4.2) des recettes du péage de 2006 à 2008.

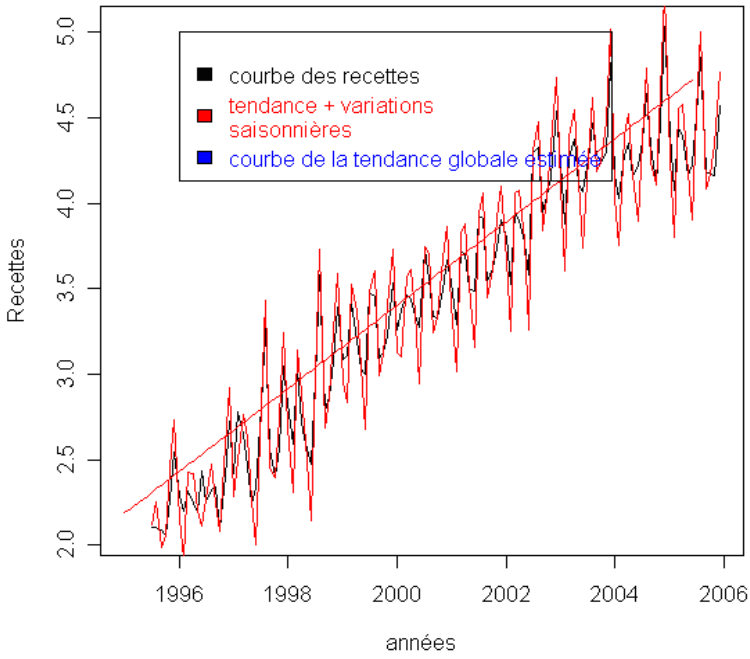


FIG. 4.3 – Graphe de la série corrigée des variations saisonnières.

4.1.3 Modélisation des recettes mensuelles du péage routier par un processus ARIMA

Nous adoptons ici une démarche différente de la première,il s'agit de celle dans [1] par Box & Jenkins pour modéliser une série chronologique présentant de fortes variations saisonnières comme la nôtre (figure 4.1). Nous présentons cette démarche en l'appliquant à la chronique de nos recettes, qui n'est en fait qu'une série chronologique s'étendant sur 11 années, dans laquelle la fréquence des observations est mensuelle et la périodicité est annuelle. On supposera dans toute cette étude que cette période est de 12 mois. Cette démarche consiste à :

1. Premièrement, éliminer les variations saisonnières de la chronique par standardisation en se servant de la transformation

$$Z_{\nu,\tau} = \frac{X_{\nu,\tau} - \mu_\tau}{\sigma_\tau},$$

où $\nu = 1, 2, \ldots, 11$; $\tau = 1, 2, \ldots, 12$; ν est le numéro d'une année ; τ est le numéro d'un mois dans une année ; $X_{\nu,\tau}$ est la recette du mois numéro τ de l'année numéro ν ; μ_τ est la moyenne empirique des recettes du mois numéro τ sur les 11 années d'observations, c'est-à-dire :

$$\mu_\tau = \frac{1}{11}\sum_{\nu=1}^{11} X_{\nu,\tau} \ ;$$

σ_τ^2 est la variance empirique des débits du mois numéro τ sur les 11 années d'observations, c'est-à-dire

$$\sigma_\tau^2 = \frac{1}{10}\sum_{\nu=1}^{10}(X_{\nu,\tau} - \mu_\tau)^2.$$

2. Deuxièmement, modéliser la série standardisée, $Z_{\nu,\tau}$ par un processus stochastique de type $ARIMA(p,d,q)$. La figure 4.4 présente l'évolution des recettes standardisée, $Z_{\nu,\tau}$. (Voir commandes C.41 en annexe pour son obtention).

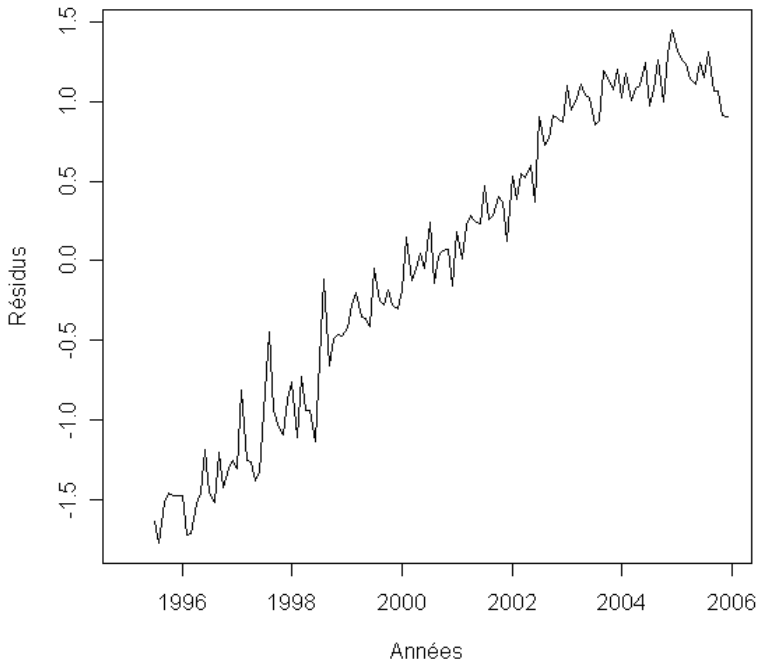


FIG. 4.4 – Evolution des recettes standardisées.

Pour la modélisation du processus $Z_{\nu,\tau}$, nous suivrons la démarche présentée dans la section I-4 du chapitre 3. En vu d'identifier l'ordre de différentiation d, observons l'autocorrélogramme (graphe de l'ACF) de la série $Z_{\nu,\tau}$ (figure 4.5). (Voir commande C.42 en annexe pour son obtention).

Ce corrélogramme montre une décroissance lente de la fonction d'auto corrélation. On peut donc penser que le processus $Z_\tau = Z_{12\,(\nu-1)+\tau \equiv Z_{\nu,\tau}}$ n'est pas stationnaire.

Observons alors l'auto-corrélogramme du processus $Y_t = (1-B)\,Z_t$ sur la figure 4.5.(Voir commande C.44 en annexe pour son obtention).

Ce corrélogramme montre que la fonction d'autocorrélation décline rapidement vers 0. On peut donc penser que le processus Y_t est stationnaire (une décroissance rapide vers 0 de la fonction d'auto corrélation est une méthode pratique d'identification des processus stationnaires, confère [7]). La probabilité critique du test de non stationnarité

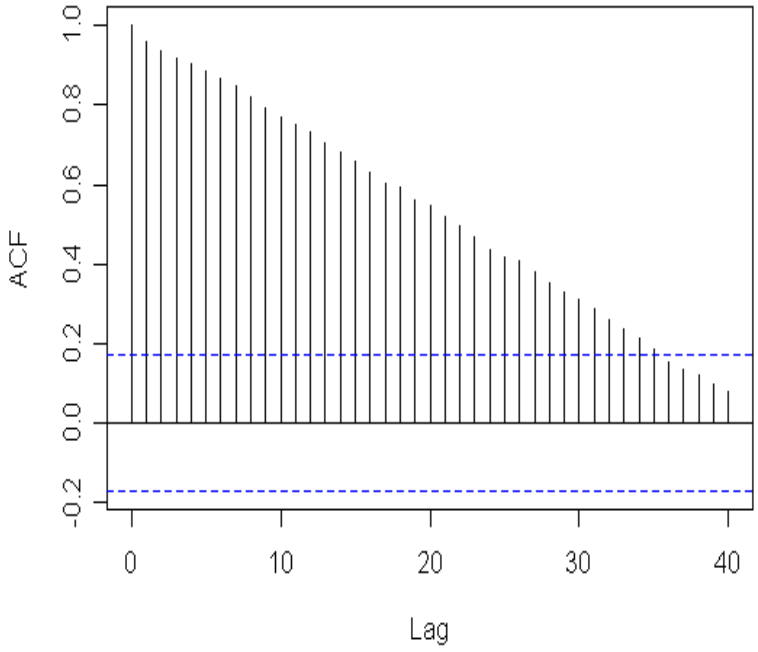


FIG. 4.5 – Corrélogramme des recettes standardisées.

de Phillips-Perron est 0.01 inférieure à 0.05, on rejette alors l'hypothèse de non stationnarité du processus Y_t avec un risque de se tromper de 5%. (Voir commande C.43 en annexe pour la réalisation de ce test). Nous retenons alors $d = 1$ comme l'ordre de différentiation du processus.

Rappelons ici quelques résultats pratiques ([7]), nous permettant d'apporter une justification supplémentaire pour le choix du nombre 1 comme ordre de différentiation du processus Z_t.

1. Les observations d'un processus stationnaire fluctuent autour d'une valeur moyenne.
2. Si l'auto corrélation de décalage 1 est égale à 0 ou est négative, le processus n'a pas besoin d'être différencié. Si l'autocorrélation de décalage 1 est inférieure à -0.5, le processus est sur différencié.

Vérification graphique de ces deux résultats.

On voit clairement sur la figure 4.7 que l'autocorrélation de décalage 1 n'est pas significative. En effet, cette valeur est à l'intérieur de la région de confiance au niveau 95% délimitée par la bande en pointillé bleue.

La courbe montrant l'évolution du processus Y_t est présentée sur la figure 4.6 ci-dessous, (voir commande C.45 en annexe pour son obtention). On voit clairement sur cette figure que les observations de ce processus fluctuent autour de la valeur moyenne 0.

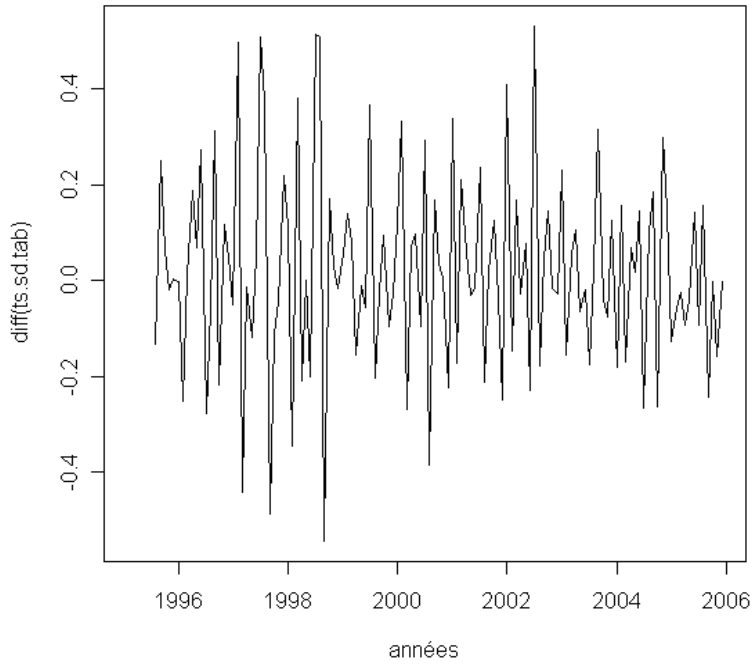


FIG. 4.6 – Evolution des recettes standardisées et différenciées.

Considérons le processus $X_t = (1-B)^2 Z_t = (1-B) Y_t$. Son autocorrélogramme est présenté sur la figure 4.7 ci-dessous, (voir commande C.46 en annexe pour son obtention). On voit clairement sur cette figure que l'autocorrélation de décalage 1 est inférieure à -0.5. Ce qui traduit une sur-différentiation du processus Z_t .

Les résultats précédents confirment le choix de la valeur 1 comme ordre de différentiation du processus Z_t. En vu de déterminer les ordres p et q, nous étudions le processus Y_t tout en supposant qu'il est un $ARMA(p,q)$. Pour cela, nous procédons par une approche empirique qui consiste à déterminer un couple (P,Q) tel que Y_t est approximativement un $AR(P)$ et un $MA(Q)$. Puis, à essayer de poursuivre la modélisation avec tous les couples (p,q) majorés par (P,Q).

Nous choisissons $Q = 22$. En effet, c'est le rang (compté à partir de 0) après lequel les autocorrélations estimées du processus Y_t sont non significatives au seuil de 5% (sur la figure 4.6, les autocorrélations estimées dont les rangs sont supérieurs à 22 sont quasiment toutes à l'intérieur de la bande en pointillé bleue). Cette bande définit une région de confiance au niveau 95%, commune à tous les estimateurs des auto corrélations. Pour le choix de P, observons la courbe de l'autocorrélogramme partielle (graphe de la fonction d'autocorrélation partielle) du processus Y_t (figure 4.8). (Voir commande C.47 en annexe pour son obtention).

Nous choisissons $p = 5$. En effet, c'est le rang (compté à partir de 1) au-delà duquel les autocorrélations partielles estimées du processus Y_t sont non significatives au seuil de 5% (sur la figure 4.8, les autocorrélations partielles estimées dont les rangs

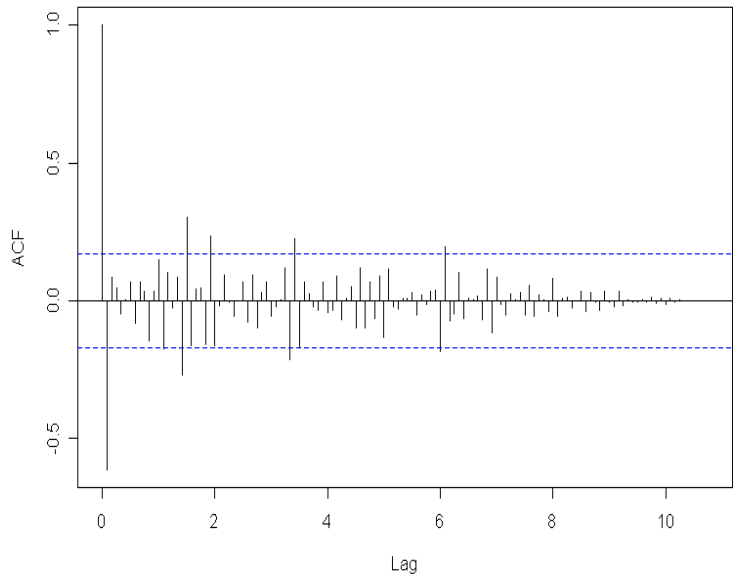


FIG. 4.7 – ACF des recettes standardisées et différentiées deux fois.

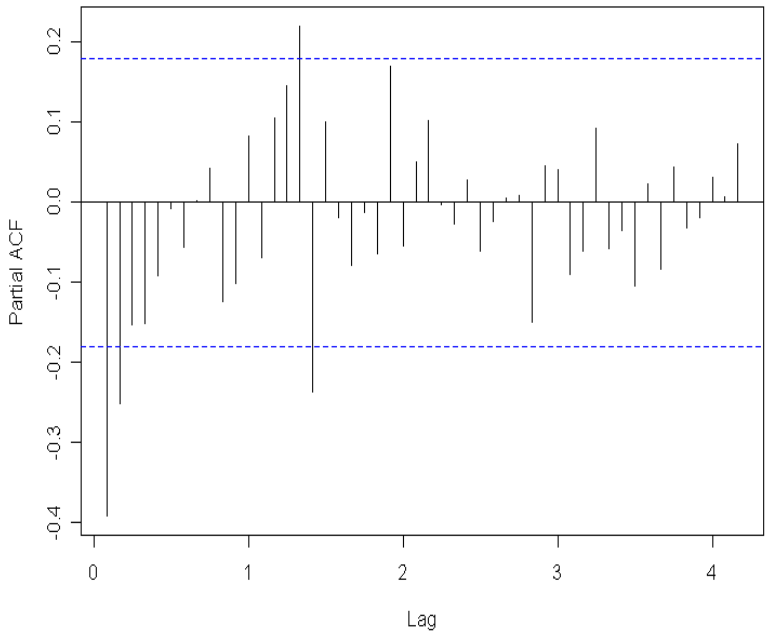


FIG. 4.8 – PACF des recettes standardisées et différentiées.

sont supérieurs à sont toutes à l'intérieur de la bande en pointillé bleue). Cette bande définit une région de confiance au niveau 95%, commune à tous les estimateurs des autocorrélations partielles.

Après de nombreuses modélisations, nous nous sommes rendus compte que la va-

riance estimée des résidus pour le modèle $ARIMA(5,1,22)$ est la plus petite parmi tous les modèles ayant rendu possibles l'estimation du modèle $ARIMA(p,1,q)$ par le logiciel R. Le modèle $ARIMA(5,1,22)$, défini par le processus Z_t a pour équation :

$$(1-B)\left(1-\sum_{i=1}^{5} ar_i\, B\right) Z_t = \left(1+\sum_{j=1}^{22} ma_j\, B^j\right)\, \varepsilon_t,$$

où ε_t est un bruit blanc de variance σ^2 ; est l'opérateur retard d'ordre p.

Les valeurs estimées des paramètres ar_i, ma_j, et σ^2 ainsi que leurs intervalles de confiances respectifs sont stockés dans le tableau.4.4. Ce tableau est obtenu en utilisant la commande C.48 disponible en annexe.

Dans le tableau 4.4 les coefficients significatifs sont ceux marqués d'un astérisque. Nous retenons ainsi le modèle ci-après :

$$(1-B)\,(1+0.0840\,B^4)\,Z_t = (1-0.3851\,B^3-0.4848\,B^5+0.4640\,B^{12}-0.6540\,B^{15}+0.7081\,B^{18})\,\varepsilon_t,$$

où ε_t est un bruit blanc de variance estimée $\widehat{\sigma}^2 = 0.020$.

Vérifions l'hypothèse de bruit blanc des résidus de ce modèle. On observe sur les graphes de la figure 4.9 obtenue en utilisant la commande C.50 que ces résidus forment effectivement un bruit blanc. permet alors de ne pas rejeter l'hypothèse nulle au seuil 5%. Nous pouvons dont conclure que ces résidus forment effectivement un bruit blanc.

Conclusion de cette modélisation

Nous terminons cette partie en disant que l'équation finale du modèle vérifié par la chronique de nos recettes mensuelles $X_{\nu,\tau}$ est :

$$X_{\nu,\tau} = \sigma_\tau\, Z_{\nu,\tau} + \mu_\tau,$$

où $Z_{\nu,\tau}$ est un processus qui suit le modèle $ARIMA(5,1,18)$ dont les paramètres sont stockés dans le tableau 4.5. Nous utiliserons cette équation sous la forme ci-dessus pour estimer les valeurs de la chronique de ces recettes.

$$\widehat{X}_{\nu,\tau} = \sigma_\tau\, \widehat{Z}_{\nu,\tau},$$

où $\nu = 1,\ldots,11$; $\tau = 1,\ldots,12$; les $\widehat{Z}_{\nu,\tau}$ les sont les valeurs du processus $Z_{\nu,\tau}$ estimés par le modèle $ARIMA(5,1,18)$. Nous présentons sur la figure 4.10 obtenue en utilisant la commande C.51, l'ajustement de ce modèle à la chronique des recettes réelles.

Afin de mieux prédire nos recettes, nous allons introduire dans notre étude d'autres variables de notre économie qui pourraient permettre d'expliquer l'évolution des recettes. Ces variables n'étant pas répertoriées à partir du même instant, le tableau contenant toutes ces données présentera plusieurs données manquantes.

Coefficients		2,50%	97,50%
ar1	-0.1588	-0,5455	0,2279
ar2	-0.0989	-0,6527	0,4549
ar3	0.3208	-0,0076	0,6493
ar4*	-0.0840	-0,5326	-0,3647
ar5*	0.6012	0.2571	0,9452
ma1	-0.3134	-0.7679	0,1412
ma2	-0.1166	-0.6874	0,4541
ma3*	-0.3851	-0,7570	-0.0132
ma4	0.2376	-0.4025	0.8771
ma5*	-0.4848	-0,8493	-0.1202
ma6	0.2691	-0.0082	0,5299
ma7	-0.0227	-0.3142	0,2688
ma8	-0.0473	-0.3711	0,2765
ma9	0.0709	-0.1541	0.2959
ma10	-0.2014	-0.4657	0.0630
ma11	0.1975	-0,0306	0.4256
ma12*	0.4640	0.2246	0.7033
ma13	-0.0271	-0.3510	0.2969
ma14	0.1078	-0.3318	0.5474
ma15	-0.1916	-0,4663	0.0831
ma16	-0.1329	-0,4644	0.1986
ma17*	-0.6540	-1.0157	-0.2922
ma18*	0.7081	0,3151	1.1010
ma19	-0.1502	-0,5483	0.2480
ma20	0.1761	-0,3446	0.6967
ma21	-0.2022	-0.5722	0.1679
ma22	0.2696	-0.0986	0.6378

TAB. 4.4 – Paramètres estimés du modèle $ARIMA(5, 1, 22)$.

4.2 Imputation des données manquantes

Vu que le taux de nos données manquantes est faible et qu'il n'y a pas de dépendance entre avoir une valeur manquante sur une variable et les autres variables du tableau de données, se référant à [5], nous utiliserons l'imputation par la moyenne qui est la plus efficace dans ce cas.

La commande C.52 nous permet d'imputer toutes les données manquantes toutes

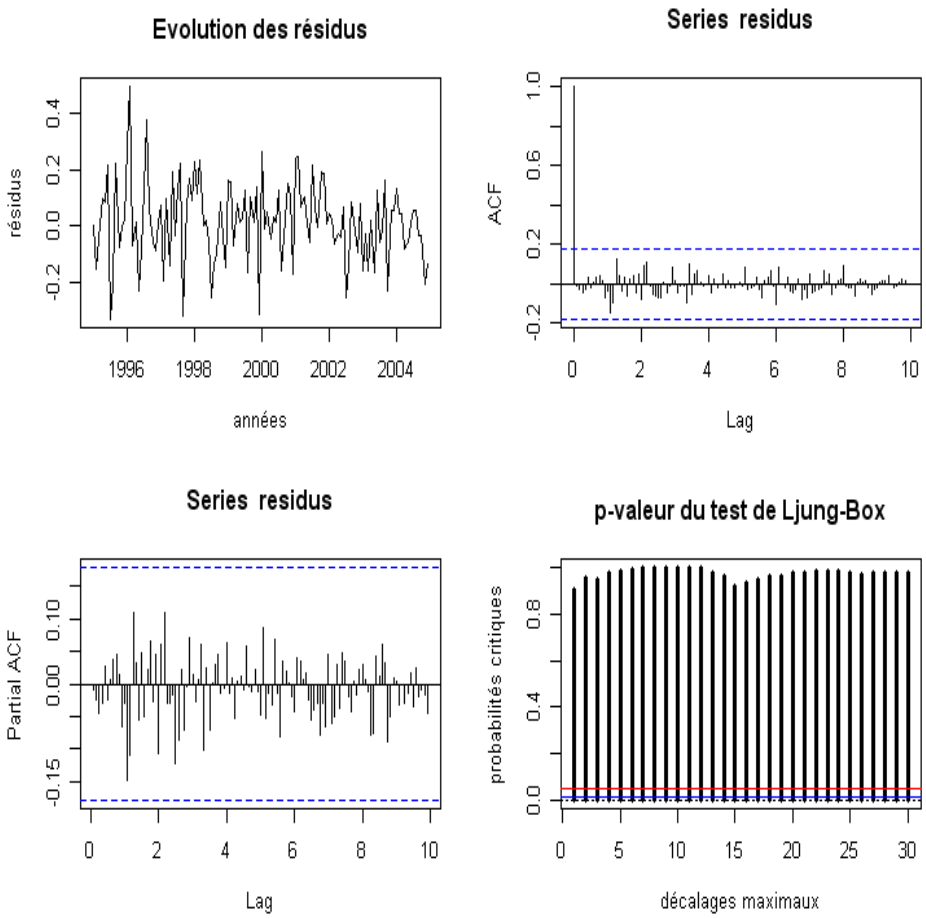


FIG. 4.9 – Diagnostic des résidus ε_t du modèle $ARIMA(5, 1, 18)$.

les données manquantes.

Nos données manquantes imputées, nous allons faire nos prédictions à l'aide d'une régression linéaire multiple.

4.3 Modélisation et prévision des recettes mensuelles du peage en fonction d'autres paramètres.

A partir de la méthode exposée dans [2], nous avons par les commandes C.53 les résultats suivants qui vont nous permettre de vérifier rapidement l'allure raisonnablement symétrique des distributions et la présence de quelques points atypiques. Les variables étant d'ordres de grandeur très différents, nous passerons au logarithme sur toutes les variables.

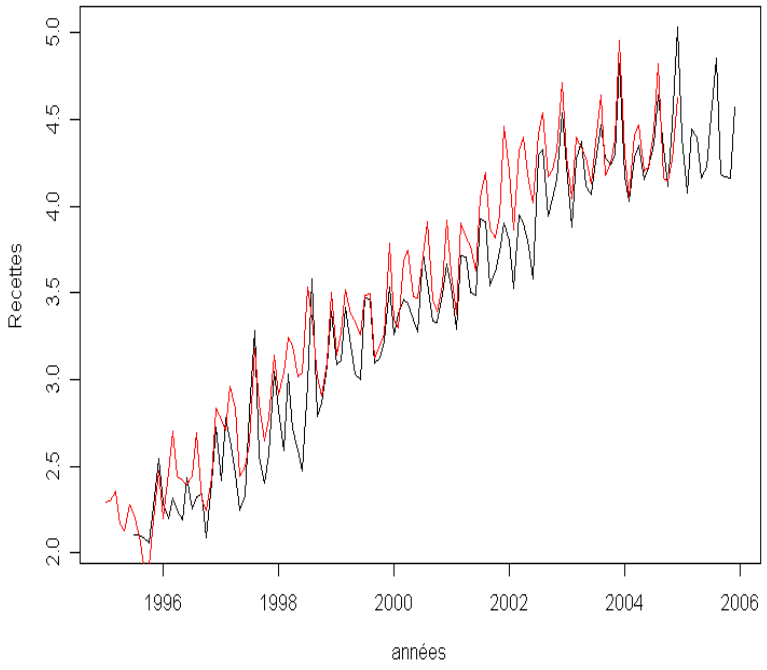


FIG. 4.10 – Ajustement de la série des résidus.

	Recette	TMI	VPAC	PSP	PGP	PMC
Minimum	19.14	13.98	11.81	5.79	5.501	5.695
1^{er} Quantile	19.46	14.08	11.88	5.94	5.687	5.822
Médiane	19.66	14.12	12.00	5.075	5.864	5.981
Moyenne	19.63	14.12	12.07	6.096	5.828	5.940
3^{me} Quantile	19.84	14.16	12.27	6.096	5.951	6.023
Maximum	20.04	14.27	12.41	6.292	6.194	6.244

TAB. 4.5 – Résumé des données de notre étude.

4.3.1 Choix de modèle " à la main " par élimination

Il est nécessaire de savoir se " débrouiller " avec les outils plus limités afin de comprendre comment fonctionnent les algorithmes de sélection automatique que propose R.

Itérer la procédure suivante :

1. Estimer et choisir, parmi les variables explicatives, celle X^j pour laquelle le test de Student $H_0 : b_j = 0$ est le moins significatif, c'est-à-dire avec la plus grande p-value. On constate que la variable la moins significative est PMC (commande C.54).

   item La retirer du modèle et recalculer l'estimation.

2. Arrêter le processus lorsque tous les coefficients sont considérés comme significati-

vement différents de 0 sauf celui du terme constant (intercept) qui reste constant. Comme c'était déjà le cas à la deuxième étape, nous avons dons comme variables explicatives : TMI, VPAC, PSP, PGP. Le modèle correspondant est :

$$(E_1) \qquad y = TMI^{1.07} \times VPAC^{0.75} \times PSP^{1.52} \times PGP^{-0.94}$$

qui a pour R^2 ajusté : 0.7922 Par les procédures automatiques identiques descendante et mixte nous obtenons les mêmes variables explicatives (commandes C.55, C.57)et les modèles correspondants sont identiques au modèle ($E1$) avec le même ajusté. Par la méthode ascendante (commande C.56), on trouve comme modèle :

$$(E_2) \qquad TMI^{1.07} \times VPAC^{0.75} \times PSP^{1.69}$$

Avec pour R^2 ajusté : 0.7923.

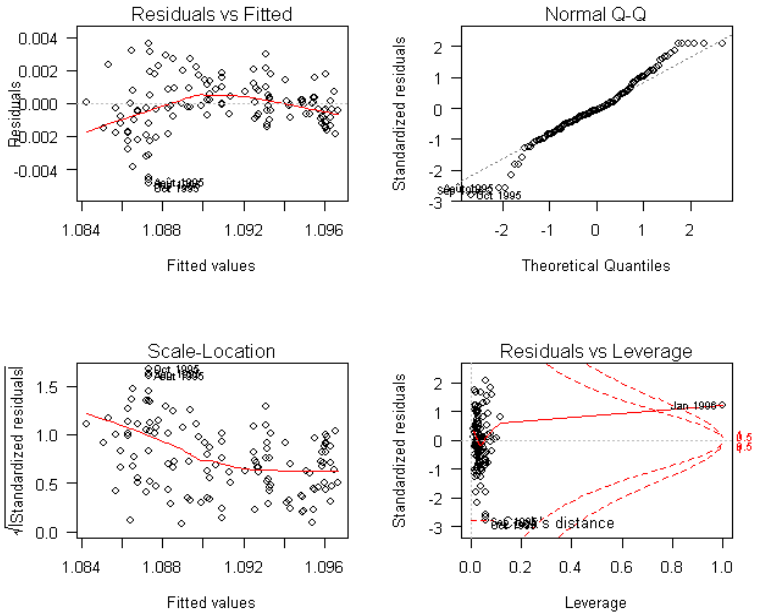


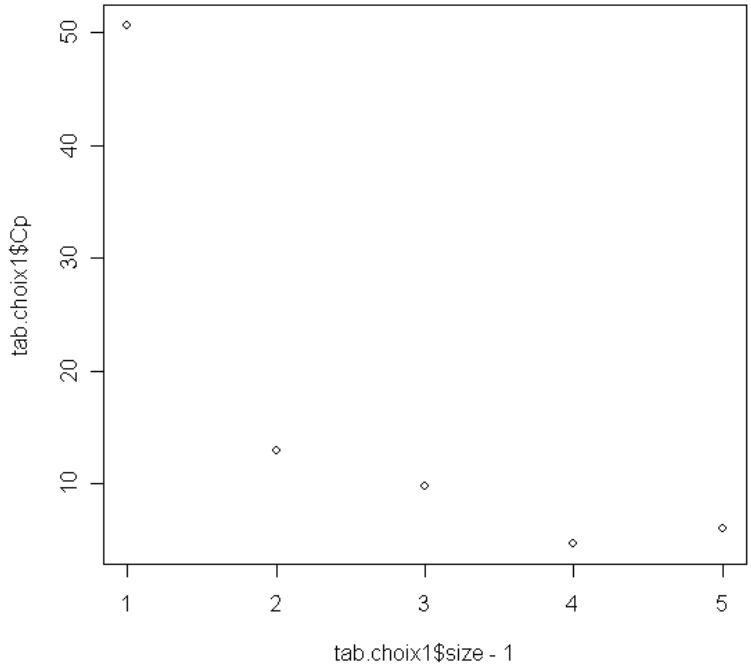
FIG. 4.11 – Diagnostics d'influence des résidus du modèle par élimination.

4.3.2 Sélection automatique du modèle

Parmi les différents algorithmes disponibles dans R et les différents critères de choix, une des façons les plus efficaces est la recherche exhaustive du meilleur modèle parmi tous les sous-modèles possibles selon l'algorithme de Furnival et Wilson. Seul le meilleur pour chaque niveau, c'est-à-dire pour chaque valeur q du nombre de variables explicatives sont donnés. Il est alors facile de choisir celui minimisant l'un des critères globaux

(C_p, R^2 ajusté,...) estimant un risque pénalisé. Cet algorithme est disponible dans le package leaps. **Meilleur modèle au sens** du C_p.

Le meilleur modèle est celui dont le C_p est minimum. Les différents C_p sont les suivants : 50.593242 ; 12.916773 ; 9.765036 ; 4.705674 ; 6.000000 (commande C.55) leur représentation graphique (commande C.58) est :



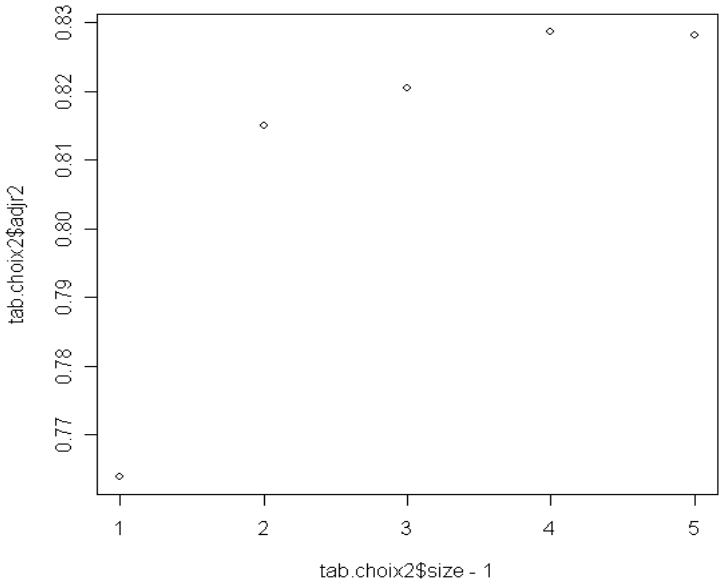
Le meilleur modèle a pour variables explicatives TMI, VPAC, PSP, PGP (voir commande C.58 en annexe). **Meilleur modèle au sens du R^2 ajusté.** Le meilleur modèle est celui dont le R^2 est maximum. Les différents R^2 sont les suivants : 0.7639508 ; 0.8149685 ; 0.8204384 ; 0.8285754 ; 0.8281772.(commande C.58). Leur représentation graphique est :

Au sens du R^2 ajusté, le meilleur modèle a également pour variables explicatives : TMI, VPAC, PSP, PGP.

Dimension	R^2	C_p	
1	0.7639508	50.59324	VPAC
2	0.8149685	12.91677	TMI VPAC
3	0.8204384	9.765036	TMI VPAC PSP PGP
4	0.8285754	4.705674	TMI VPAC PSP PGP
5	0.8281772	6	TMI VPAC PSP PGP PMC

TAB. 4.6 – Recherche du meilleur modèle au sens du du C_p et du R^2.

On note que le modèle qui en même temps minimise le C_p et maximise le R^2 au mieux est celui qui à pour variables explicatives TMI, VPAC, PSP et PGP.



4.3.3 Dernières estimations

Puisque les différents modèles : complet, meilleur C_p, meilleur R^2 ajusté sont identiques, les différents diagnostics d'influence des résidus sont les mêmes (commande C.58).

Conclusion de cette modélisation

Nous pouvons conclure après ces différentes modélisations que les variables explicatives de nos recettes sont : TMI, VPAC, PSP et PGP et le modèle que nous retiendrons ; qui correspond au plus grand ajusté et par ailleurs au plus grand C_p, est :

$$y = TMI^{1.07} \times VPAC^{0.75} \times PSP^{1.52} \times PGP^{-0.94}.$$

Afin d'évaluer la fiabilité de ce modèle, nous allons à l'aide de la commande C.59 superposer sur le même graphique les recettes réelles et les recettes prédites (Figure 4.13)

Ce deuxième modèle peut avoir des intérêts de prévision à très court terme (échéance 1 mois) des recettes mensuelles du péage. En effet les différentes variables explicatives de notre modèle s'observent chaque mois avant les recettes : le trafic moyen imposable et les prix des carburants pour un mois quelconque sont connus au plus tard au début de ce mois, les données sur le volume du parc automobile quant à elles sont annuelles alors que les données sur les recettes sont disponibles au plus tôt le 1er du mois suivant.

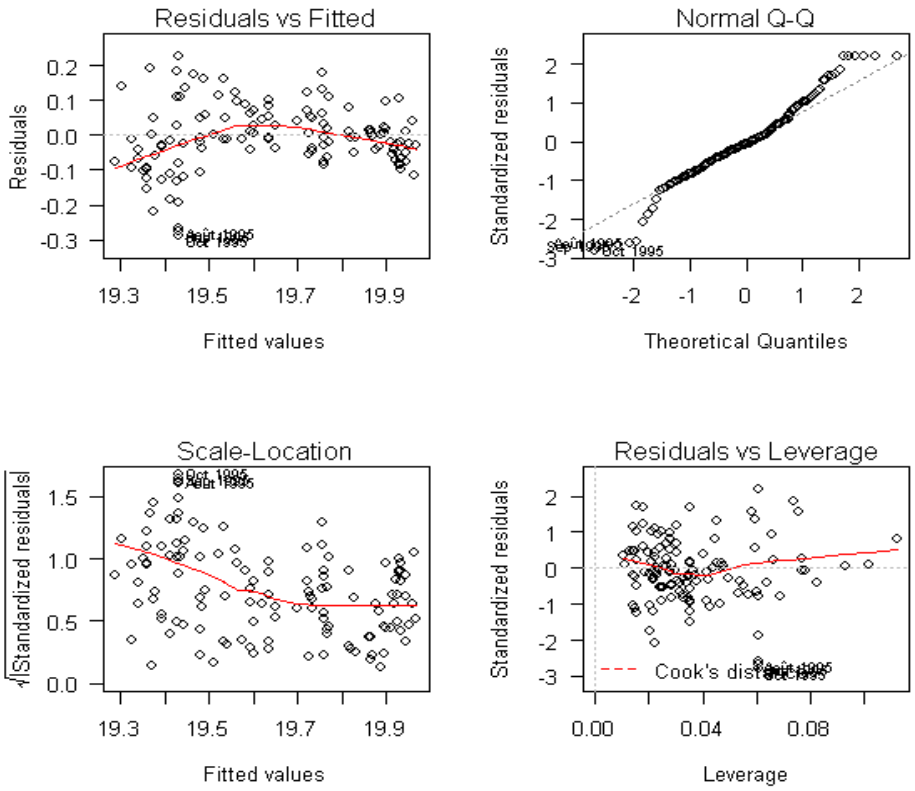


FIG. 4.12 – Diagnostics d'influence des résidus.

4.4 Erreurs quadratiques d'ajustement

Nous présentons dans le tableau 4.6 obtenu en utilisant la commande C.60, les erreurs quadratiques liées aux différents ajustements.

	Nombre de mois	er (%)	em (FCFA)
Tendance générale	132	6.63	122920236
Prévisions corrigées	132	5.76	13752385
Prévisions à base d'un modèle $ARIMA$	132	2.93	10147934
Modèle économétriques	132	3.01×10^{-8}	.01042857

TAB. 4.7 – Erreurs quadratiques d'ajustement

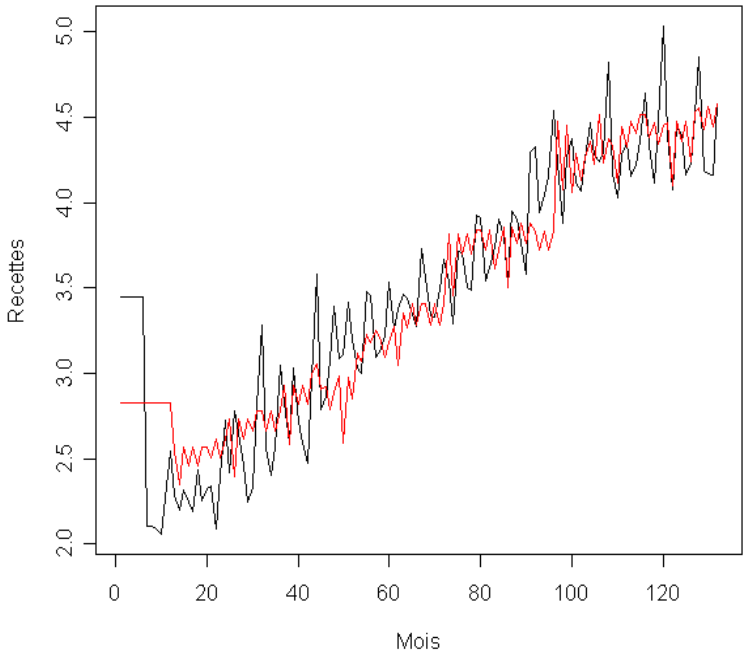


FIG. 4.13 – Evolution des recettes réelles et des recettes prédictes.

Chapitre 5

Conclusion générale

Dans le but de reverser la somme qui lui incombe dans l'entretien et la construction du réseau routier au Cameroun, le péage routier se doit de mettre sur pied des stratégies d'optimisation de ces recettes.

Pour le faire, nous avons essayé au préalable de faire des prévisions temporelles des recettes futures sur la base des recettes passées uniquement en nous inspirant premièrement de la tendance générale. Ensuite nous avons fait une prévision corrigée des variations saisonnières et enfin nous avons utilisé un processus de type $ARIMA$. Nous avons ensuite fait une prévision prenant en compte d'autres variables de notre économie afin de prédire les recettes à venir en tenant compte du TMI, du VPAC, du PSP et du PGP. Les prévisions temporelles ont été jugées plus satisfaisantes pour des prévisions à moyen et à long terme, alors que la régression linéaire multiple est adaptée pour les prévisions à court terme. Aussi ce dernier modèle a été jugé plus satisfaisant que les précédents en termes d'erreurs quadratiques d'ajustement. Par ailleurs, des modèles de simulation ont été mis au point.

Notons cependant que cette étude aurait pu être faite à l'aide des séries chronologiques multivariées, ceci pour une meilleure consistance des résultats vu que notre régression ne tient pas compte de l'évolution des données dans le temps.

Chapitre 6

Annexes : Programmes R utilisés

Elle est composée de deux parties :

- La présentation des différents tableaux
- Commandes et programmes R utilisées

Nous y présentons les différents tableaux et programmes qui ont permis à la réalisation de notre étude.

6.1 Commandes et programmes R utilisés

Lecture des données dans la console du logiciel R

Les données sont dans les tableaux 2.1 à 2.5. Ces tableaux sont enregistrés dans l'ordinateur, comme classeurs d'Excel dans des fichiers nommés : " Recettesgrales ", " Recettes 2003 ", " Recettes 2004 ", " Recettes 2005 ", " Recettes 2006 " et " tableau ". Ces tableaux sont importés sous R à l'aide des commandes.

```
library(xlsReadWrite) #Afin que les tableaux puissent être lus
tels quels dans R sans être enregistrés sous extension txt
tab1=read.xls("Recettesgrales") ; tab3=read.xls("Recettes 2003") ;
tab4=read.xls("Recettes 2004") ; tab5=read.xls("Recettes 2005") ;
tab6=read.xls("Recettes 2006") ;
tab=read.xls("tableau",colClasses="numeric",rowNames=T) ;
```

C.1 : Obtention de la figure 2.1

Nous convertissons la tableau tab1 en une série temporelle univariée en utilisant la commande

```
tstab1=ts(as.vector(as.matrix(tab1)),start=c(1995,1),end=c(2007,11),frequency=12
```

La fonction suivante permet de tracer la courbe d'évolution d'une série temporelle tout en faisant ressortir éventuellement les différentes périodes à l'intérieur des bandes verticales.

```
eda.ts <- function (x, bands=FALSE) # x est un vecteur ou une
série temporelle univariée. { if(!is.ts(x)) x <- ts(x)
plot(x,xlab="années",ylab="Recettes") if(bands) { a <- time(x) i1
<- floor(min(a)) i2 <- ceiling(max(a)) y1 <- par('usr')[3] y2 <-
par('usr')[4] if( par("ylog") ) { y1 <- 10^y1 y2 <- 10^y2 } for (i
in seq(from=i1, to=i2-1, by=2)) { polygon( c(i,i+1,i+1,i),
c(y1,y1,y2,y2), col='grey', border=NA ) } lines(x) } }
```

Nous obtenons finalement la figure 2.1 en appliquant successivement les commandes

```
eda.ts(tstab1,bands=TRUE) title(main=list("FIG.2.1 Evolution des
Recettes du péage routier Camerounais", col=4, cex=1, font=2))
```

C.2 Obtention de la figure 2.2

Nous utilisons successivement les commandes

```
mattab1=as.matrix (tab1);
matplot(mattab1[,1:4],type="l",xlab="Mois",ylab="Recettes(FCFA)",
ylim=range(mattab1, na.rm=TRUE)); legend(x=8,y=500000000,c("A
1995","A 1996","A 1997","A 1998"),col=1:4, fill=1:4, text.col=1:4)
; title(main=list("FIG.2.2 Evolution Mensuelle des Recettes du
péage routier Cameroun 1995 à 1998",col=4,cex=0.75,font=2));
```

C.3 Obtention de la figure 2.3

Nous utilisons les commandes successives

```
matplot(mattab1[,5:8],type="l",xlab="Mois",ylab="Recettes(FCFA)",ylim=range(matt
na.rm=TRUE)); legend(x=9,y=500000000,c("A 1999","A 2000","A
2001","A 2002"),   col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.2.3 Evolution Mensuelle des Recettes du péage
routier cameroun 1998 à 2002", col=4,cex=0.75,font=2));
```

C.4 Obtention de la figure 2.4

```
matplot(mattab1[,9:13],type="l",xlab="Mois",ylab="Recettes(FCFA)",
ylim=range(mattab1,  na.rm=TRUE)); legend(x=3,y=300000000,c("A
2003","A 2004","A 2005","A 2006","A 2007"),col=1:5,fill=1:5,text.col=1:5);
 title(main=list("FIG.2.4 Evolution
Mensuelle des Recettes du péage routier camerounais de 2003 à
2007", col=4, cex=0.75,font=2));
```

C.5 Obtention de la figure 2.5

```
mattab3=as.matrix (tab3);
matplot(mattab3[,1:5],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=47000000,c("P  EDEA","P NKOMETOU","P MBANGA","P
TIKO", "P BOUMNYEBEL"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.5 Evolution Mensuelle des Recettes des
postes de péage les plus fréquentés du Cameroun en 2003",
col=4,cex=0.75,font=2));
```

C.6 Obtention de la figure 2.6

```
matplot(mattab3[,6:10],type="l",xlab="Mois",
ylab="Recettes(FCFA)"); legend(x=2,y=30000000,c("P  MBANKOMO","P
MANJO","P NSIMALEN","P BANDJA","P FOUMBOT"),
col=1:5,fill=1:5,text.col=1:5); title(main=list("FIG.2.6 Evolution
Mensuelle des Recettes des postes de MBANKOMO-
MANJO-NSIMALEN-BANDJA-FOUMBOT en 2003", col=4,cex=0.75,font=2));
```

C.7 Obtention de la figure 2.7

```
matplot(mattab3[,11:15],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=17500000,c("P BAYANGAM","P BAFIA","P MATAZEM","P
DSCHANG","P AWAE"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.7 Evolution Mensuelle des Recettes des
postes de BAYANGAM-BAFIA-MATAZEM-DSCHANG-AWAE en
2003",col=4,cex=0.75,font=2));
```

C.8 Obtention de la figure 2.8

```
matplot(mattab3[,16:20],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=7000000,c("P MAGADA ","P BAMENA","P KAREWA","P
MENGONG","P MEME"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.8 Evolution Mensuelle des postes de
MAGADA-BAMENA-KAREWA-MENGONG-MEME en
2003",col=4,cex=0.75,font=2));
```

C.9 Obtention de la figure 2.9

```
matplot(mattab3[,21:25],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=9,y=5500000,c("P KRIBI","P NLOUP","P DJABI","P
```

```
TCHABAL","P BOUAM"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.9 Evolution Mensuelle des Recettes des
postes de KRIBI-NLOUP- DJABI-TCHABAL-BOUAM en 2003",
col=4,cex=0.75,font=2));
```

C.10 Obtention de la figure 2.10

```
matplot(mattab3[,26:30],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=3000000,c(" P WAZA","P KOUSSERI","P ESSONGO",
          "P NKOLOTOUTOU","P GAZAWA"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.10 Evolution Mensuelle des Recettes des
postes de WAZA-KOUSSERI-ESSONGO-NKOLOTOUTOU-GAZAWA en 2003",
col=4,cex=0.75,font=2));
```

C.11 Obtention de la figure 2.11

```
matplot(mattab3[,31:34],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=8,y=1000000,c("P MAYO.OULO","P YAGOUA","P NGATT ","P
BEKASICHI"), col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.2.11 Evolution Mensuelle des Recettes  des
postes de péage les moins fréquentés au cameroun en
2003",col=4,cex=0.75,font=2));
```

C.12 Obtention de la figure 2.12

```
mattab4=as.matrix (tab4);
matplot(mattab4[,1:5],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=48000000,c("P  NKOMETOU","P EDEA","P MBANGA",
          "P MBANKOMO","P BOUMNYEBEL"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.12 Evolution Mensuelle des Recettes des
postes de péage les plus fréquentés du Cameroun en 2004", col=4,
cex=0.75,font=2));
```

C.13 Obtention de la figure 2.13

```
matplot(mattab4[,6:10],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=7,y=43000000,c("P  MANJO","P TIKO","P NSIMALEN","P
BANDJA", "P BAYANGAM"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.13 Evolution Mensuelle des Recettes des
postes de TIKO-MANJO- NSIMALEN-BANDJA-BAYANGAM en
```

```
2004",col=4,cex=0.75,font=2));
```

C.14 Obtention de la figure 2.14

```
matplot(mattab4[,11:15],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=17000000,c("P FOUMBOT","P MATAZEM","P BAFIA","P
AWAE", "P DSCHANG"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.14 Evolution Mensuelle des Recettes des
postes de FOUMBOT- MATAZEM-BAFIA-AWAE-DSCHANG en
2004",col=4,cex=0.75,font=2));
```

C.15 Obtention de la figure 2.15

```
matplot(mattab4[,16:20],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=4,y=4500000,c("P MENGONG ","P BAMENA","P KAREWA","P
MAGADA", "P MEME"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.15 Evolution Mensuelle des postes de
MENGONG-BAMENA- KAREWA-MAGADA-MEME en 2004",
col=4,cex=0.75,font=2));
```

C.16 Obtention de la figure 2.16

```
matplot(mattab4[,21:25],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=6000000,c("P NLOUP","P KRIBI","P DJABI","P
TCHABAL","P WAZA"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.16 Evolution Mensuelle des Recettes des
postes de NLOUP-KRIBI- DJABI-TCHABAL-WAZA en 2004", col=4,
cex=0.75, font=2));
```

C.17 Obtention de la figure 2.17

```
matplot(mattab4[,26:30],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=1600000,c("P BOUAM","P KOUSSERI","P ESSONGO",
         "P NKOLOTOUTOU","P GAZAWA"), col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.17 Evolution Mensuelle des Recettes des
postes de BOUAM-KOUSSERI-ESSONGO-NKOLOTOUTOU-GAZAWA en 2004
",col=4,cex=0.75,font=2));
```

C.18 Obtention de la figure 2.18

```
matplot(mattab4[,31:34],type="l",xlab="Mois",ylab="Recettes(FCFA)");
```

```
legend(x=8,y=1000000,c("P MAYO.OULO","P YAGOUA","P NGATT ","P
BEKASICHI"), col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.2.18 Evolution Mensuelle des Recettes  des
postes de péage les moins fréquentés au cameroun en
2004",col=4,cex=0.75,font=2));
```

C.19 Obtention de la figure 2.19

```
mattab5=as.matrix (tab5);
matplot(mattab5[,1:5],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=48000000,c("P  EDEA","P NKOMETOU","P MBANGA","P
TIKO", "P MBANKOMO"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.19 Evolution Mensuelle des Recettes des
postes de péage les plus fréquentés du Cameroun en
2005",col=4,cex=0.75,font=2));
```

C.20 Obtention de la figure 2.20

```
matplot(mattab5[,6:10],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=38000000,c("P BOUMNYEBEL","P MANJO","P NSIMALEN",
           "P BANDJA","P FOUMBOT"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.20 Evolution Mensuelle des Recettes des
postes de BOUMNYEBEL-MANJO-NSIMALEN-BANDJA-FOUMBOT en
2005",col=4,cex=0.75,font=2));
```

C.21 Obtention de la figure 2.21

```
matplot(mattab5[,11:15],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=9,y=20000000,c("P BAYANGAM","P MATAZEM","P BAFIA","P
AWAE","P DSCHANG"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.21 Evolution Mensuelle des Recettes des
postes de BAYANGAM-MATAZEM-BAFIA-AWAE-DSCHANG en 2005", col=4,
      cex=0.75, font=2));
```

C.22 Obtention de la figure 2.22

```
matplot(mattab5[,16:20],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=7800000,c("P MENGONG","P MAGADA","P MEME","P KAREWA",
"P BAMENA"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.22 Evolution Mensuelle des postes de
```

```
MENGONG-MAGADA-MEME-KAREWA-BAMENA en
2005",col=4,cex=0.75,font=2));
```

C.23 Obtention de la figure 2.23

```
matplot(mattab5[,21:25],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=6000000,c("P DJABI","P NLOUP","P KRIBI","P TCHABAL",
          "P KOUSSERI"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.23 Evolution Mensuelle des Recettes des
postes de DJABI-NLOUP- KRIBI-TCHABAL-KOUSSERI en 2005", col=4,
cex=0.75,
     font=2));
```

C.24 Obtention de la figure 2.24

```
matplot(mattab5[,26:30],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=8,y=3800000,c("P  WAZA","P BOUAM","P ESSONGO","P
NKOLOTOUTOU", "P GAZAWA"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.24 Evolution Mensuelle des Recettes des
postes de WAZA-BOUAM-ESSONGO-NKOLOTOUTOU-GAZAWA en
2005",col=4,cex=0.75,font=2));
```

C.25 Obtention de la figure 2.25

```
matplot(mattab5[,31:34],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=8,y=1000000,c("P MAYO.OULO","P YAGOUA","P NGATT ","P
BEKASICHI"),col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.2.25 Evolution Mensuelle des Recettes  des
postes de péage les moinsfréquentés au cameroun en
2005",col=4,cex=0.75,font=2));
```

C.26 Obtention de la figure 2.26

```
mattab6=as.matrix (tab6);
matplot(mattab6[,1:5],type="l",xlab="Mois",ylab="Recettes(FCFA)")
legend(x=8,y=45000000,c("P  EDEA","P NKOMETOU","P MBANKOMO",
          "P MBANGA","P BOUMNYEBEL"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.26 Evolution Mensuelle des Recettes des
postes de péage les plus fréquentés du Cameroun en
2006",col=4,cex=0.75,font=2));
```

C.27 Obtention de la figure 2.27

```
matplot(mattab6[,6:10],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=8,y=33000000,c("P  TIKO","P MANJO","P NSIMALEN","P
BANDJA", "P BAYANGAM"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.27 Evolution Mensuelle des Recettes des
postes de TIKO-MANJO-NSIMALEN-BANDJA-BAYANGAM en
2006",col=4,cex=0.75,font=2));
```

subsection*C.28 Obtention de la figure 2.28

```
matplot(mattab6[,11:15],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=2,y=5000000,c("P FOUMBOT","P MATAZEM","P BAFIA","P AWAE",
"P DSCHANG"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.28 Evolution Mensuelle des Recettes des
postes de FOUMBOT-MATAZEM-BAFIA-AWAE-DSCHANG en
2006",col=4,cex=0.75,font=2));
```

C.29 Obtention de la figure 2.29

```
matplot(mattab6[,16:20],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=4,y=5500000,c("P KAREWA ","P MAGADA","P MENGONG","P
MEME", "P DJABI"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.29 Evolution Mensuelle des postes de
KAREWA-MAGADA-MENGONG-MEME-DJABI en 2006",col=4,cex=0.75,font=2));
```

C.30 Obtention de la figure 2.30

```
matplot(mattab6[,21:25],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=6,y=3000000,c("P BAMENA","P KRIBI","P NLOUP","P BOUAM",
          "P KOUSSERI"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.30 Evolution Mensuelle des Recettes des
postes de BAMENA-KRIBI- NLOUP-BOUAM-KOUSSERI en 2006", col=4,
cex=0.75, font=2));
```

C.31 Obtention de la figure 2.31

```
matplot(mattab6[,26:30],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=8,y=3500000,c("P TCHABAL","P WAZA","P ESSONGO",P
NKOLOTOUTOU", "P GAZAWA"),col=1:5,fill=1:5,text.col=1:5);
title(main=list("FIG.2.31 Evolution Mensuelle des Recettes des
```

```
postes de TCHABAL-WAZA- ESSONGO-NKOLOTOUTOU-GAZAWA en
2006",col=4,cex=0.75,font=2));
```

C.32 Obtention de la figure 2.32

```
matplot(mattab6[,31:34],type="l",xlab="Mois",ylab="Recettes(FCFA)");
legend(x=6,y=650000,c("P MAYO.OULO","P YAGOUA","P NGATT ","P
BEKASICHI"), col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.2.32 Evolution Mensuelle des Recettes  des
postes de péage les moins fréquentés au cameroun en
2006",col=4,cex=0.75,font=2));
```

C.33 Obtention de la figure 2.33

```
ts.tab1 =ts(as.vector(as.matrix(tab$TMI)),start=c(1995,1),frequency=12)
plot(ts.tab1)
```

C.34 Obtention de la figure 2.34

```
ts.tab2 =ts(as.vector(as.matrix(tab$VPAC)),start=c(1995,1),frequency=12)
plot(ts.tab2)
```

C.35 Obtention de la figure 2.36

```
ts.tab3 =ts(as.vector(as.matrix(tab$PSP)),start=c(1995,1),frequency=12)
plot(ts.tab3)
```

C.37 Obtention de la figure 2.37

```
ts.tab4 =ts(as.vector(as.matrix(tab$PGP)),start=c(1995,1),frequency=12)
plot(ts.tab4)
```

C.38 droite de régression des écarts-types

Nous utilisons les commandes successives

```
tab=tab[,-c(12:13)]
 ts.tab =ts(as.vector(as.matrix(tab)),start=c(1995,1),frequency=12)
#Droite de régression des écarts-types#
ecart=apply(tab,2,sd,na.rm=T) moyenne=apply(tab,2,mean,na.rm=T)
reg=lm(ecart~moyenne) r=summary(reg)
```

C.39 prévision simples

```
t=1:length(ts.tab) reg1=lm(ts.tab~t) summary(reg1)
plot(ts.tab/10^8,xlab="années",ylab="Recettes(*10^8)")
lines(ts(reg1$fitted.values/10^8,frequency=12,start=1995),col=2)
legend(x=1996,y=5,c("Courbe des recettes réelles","Courbe des
recettes prédites"), col=1:4,fill=1:4,text.col=1:4);
title(main=list("FIG.4.1 Tendance des Recettes du péage routier
camerounais ", col=4,cex=0.75,font=2));
x=data.frame(t=seq(length(ts.tab)+1,lengnth(ts.tab)+36,1))
 predict(reg1,x,interval="prediction")
```

C.40 obtention de la figure 4.2

Ce programme prend en entrée une série temporelle "serie" et l'ordre de la série moyenne mobile à calculer "p", et retourne la série des moyennes mobiles centrées.

```
moy_mob_cent<-function(serie,p) {
if(p%%2!=0)
{ cas_imp<-function(serie,p) { serie_vec<-c(serie)
n<-length(serie_vec)
m=p%/%2
k=n-2*m tab=matrix(,nrow=k,ncol=p) i=1
while(p+i-1<=n){tab[i,]=serie_vec[i:(p+i-1)];i=i+1}
serie_mob1<-apply(tab,1,mean) # vecteur des moyennes mobiles #
td<-start(serie)[2]+(p-1)/2 freq=frequency(serie)
serie_mob<-ts(serie_mob1,start=c(start(serie)[1]+td%/%freq,td%%freq),frequency=f
# Transformation en série temporelle # } t<-cas_imp(serie,p) }
else { cas_pair<-function(serie,p) { serie_vec<-c(serie)
n<-length(serie_vec)
m=p%/%2
k=n-2*m tab=matrix(,nrow=k,ncol=p+1) i=1
while(p+i<=n){tab[i,]=serie_vec[i:(p+i)];i=i+1}
tab[,c(1,p+1)]=tab[,c(1,p+1)]*0.5
serie_mob1<-apply(tab,1,function(x){a=sum(x)/p;a}) # vecteur des
moyennes mobiles # td<-start(serie)[2]+p/2 freq=frequency(serie)
serie_mob<-ts(serie_mob1,start=c(start(serie)[1]+td%/%freq,td%%freq),frequency=f
# transformation en série temporelle # } t<-cas_pair(serie,p) } t
} #=====================================================#
serie.moy=moy_mob_cent(ts.tab,p=12)
```

```
plot((ts.tab)/10^8,xlab="Années",ylab="Recettes(*10^8")
lines((serie.moy)/10^8,col=2) legend(x=1996,y=5,legend=c("courbe
des recettes","courbe des moyennes mobiles
centrées"),col=1:2,text.col=1:2,fill=1:2) title(main=list("FIG.4.2
Série des moyennes mobiles d'ordre 12",cex=1,col=4,font=2))
```

C.41 Prévisions corrigées

Ce programme prend en entrée une série temporelle "serie" et retourne les coefficients saisonniers corrigés ainsi que la série corrigée des variations saisonnières. Ceci en désaisonnalisant cette série à l'aide d'une moyenne mobile centrée d'ordre "p" et en supposant que le modèle est de type additif.

```
desaisonnalisation<-function(serie,p) { diff_sais_serie<-serie -
moy_mob_cent(serie,p) # calcul des différences saisonnières #
c1<-matrix(transform(diff_sais_serie),byrow=T,ncol=frequency(serie),
byrow=TRUE) coef_sais<-apply(c1,2,mean,na.rm=TRUE) # coefficients
saisonniers # coef_sais_corr<-coef_sais-mean(coef_sais) #
coefficients saisonniers corrigés # cvs<-serie - coef_sais_corr #
série corrigée des variations saisonnières #
sortie=list(Coefficients_Saisonniers_Corrigés=coef_sais_corr,Série_Corrigée_Des_
sortie } #=====================================================#

des=desaisonnalisation(ts.tab,p=12) des[[1]]

plot(ts.tab,xlab="années",ylab="Recettes") lines(des[[2]],col=2)
t=1:length(des[[2]]) reg2=lm(c(des[[2]])~t) abline(reg2,col=4)
legend(x=1996,y=5,legend=c("courbe des recettes","tendance + des
variations saisonnières","courbe de la tendance globale
estimée"),text.col=c(1,2,4),col=c(1,2,4),fill=c(1,2,4))
title(main=list("FIG.4.3 Graphe de la série corrigée des
variations saisonnières", col=4,cex=0.75,font=2)); summary(reg2)
x=data.frame(t=seq((length(des[[2]])+1),(length(des[[2]])+36),1))
 predict(reg2,x,interval="prediction")

\subsection*{C.42 Obtention de la figure 4.4}
m.tab=as.matrix(tab)
 sd.tab=t(scale(t(m.tab)))
 ts.sd.tab =ts(as.vector(as.matrix(sd.tab)),start=c(1995,1),frequency=12)
 plot(ts.sd.tab,xlab="Années",ylab="Résidus")
```

```
 title(main=list("FIG.4.4 Evolution des recettes standardisées",cex=1,col=4,font
```

C.43 Obtention de la figure 4.5

```
acf(c(ts.sd.tab),na.action=na.pass,lag.max=50,main="FIG 4.5
corrélogramme des recettes standardisées")
```

C.44 Test de Philips-perron 4.5

```
PP.test(diff(ts.sd.tab),lshort =F)
```

C.45 Obtention de la figure 4.6

```
acf(diff(ts.sd.tab),na.action=na.pass,main="FIG4.6 ACF des
recettes standardisées et différenciées",lag.max=150)
```

C.46 Obtention de la figure 4.7

```
plot(diff(ts.sd.tab),xlab="années") title(main=list("FIG.4.7
Evolution des recettes standardisées et
différenciées",col=4,cex=1,font=2))
```

C.47 Obtention de la figure 4.8

```
acf(diff(diff(ts.sd.tab)),na.action=na.pass,lag.max=150,main="FIG4.8
ACF des recettes standardisées et différenciées deux fois")
```

C.48 Obtention de la figure 4.9

```
pacf(diff(ts.sd.tab),na.action=na.pass,lag.max=50,main="FIG.4.9
PACF des recettes standardisées et différenciées")
```

subsection*C.49 Obtention du tableau 4.4

```
model=arima(ts.sd.tab,order=c(5,1,22),include.mean=F,method =
"CSS",transform.pars=F) confint(model)
```

C.50 Obtention du tableau 4.5

La fonction suivante prend entrée une série chronologique et trace le graphe des p-valeurs du test de Ljung-Box en fonction des ordres maximaux de décalages fixés, jusquà un certain rang éventuellement précisé.

```
plot.box.ljung <- function (z, k=15, main="p-valeur du test de
Ljung-Box") {
   p <- rep(NA, k)
   for (i in 1:k)
    {
      p[i] <- Box.test(z, i, type = "Ljung-Box")$p.value
    }
  plot(p, type='h', ylim=c(0,1), lwd=3, main=main,xlab="décalages maximaux",
  ylab="probabilités critiques")
  abline(h=c(0,.05),lty=3)
  abline(0.05,0,col=2)
  abline(0.01,0,col=4)
} #=====================================================#
 Nous effectuons successivement les commandes suivantes pour avoir les graphes d
figure 4.7

residus=model$residuals par(mfrow=c(2,2))
plot(residus,xlab="années",ylab="résidus",main="Evolution des
résidus") acf(residus,lag.max=150) pacf(residus,lag.max=150)
plot.box.ljung(residus,k=30)
```

C.51 Obtention de la figure 4.11

```
residus=model$residuals ts.rec.stand.est=ts.sd.tab-residus
ts.rec.est=(ts.rec.stand.est*sd)+moy
ts.rec=ts(as.vector(as.matrix(tab)),start=c(1995,1),frequency=12)
plot((ts.tab)/10^8,xlab="années",ylab="Recettes",main="FIG 4.11
Ajustement de la série des recettes")
lines((ts.rec.est)/10^8,col=2) legend(x=1996,y=5,legend=c("courbe
des recettes","courbe des recettes ajustées"),
text.col=1:2,fill=1:2,col=1:2)
```

C.52 Imputation des données manquantes

```
apply(tab,2,function(x){sum(is.na(x))/length(x)}) #donne le
```

```
pourcentage de
 # données manquantes pour chaque variable.
tab=as.data.frame(apply(tab,2,function(x){x[is.na(x)]=mean(x,na.rm=T);x}))
```

C.53 Estimation du modèle complet avec toutes les autres variables

Nous effectuons les commandes successives

```
tab=log(tab) ; Recettes=log(Recettes) ; TMI=log(TMI) ;
VPAC=log(VPAC) ; PSP=log(PSP) ; PGP=log(PGP) ; PMC=log(PMC)
summary(tab)# Affiche les résultats numériques#
lm.tab=lm(Recettes~.,data=tab) # Résultats numériques#
summary(lm.tab) #Regroupement des graphiques sur la même figure#
par(mfrow=c(2,2)) # Résidus et points influents#
plot(lm.tab,las=1)
```

C.54 Choix de modèle à la main

```
lm.tab2=lm(Recettes~TMI+VPAC+PSP+PGP+PMC, data=tab) anova(lm.tab2)
lm.tab2=lm(Recettes~TMI+VPAC+PSP+PGP, data=tab) summary(lm.tab2)
REG=lm.tab2
```

C.55 Choix de modèle par la méthode descendante

```
step.tab1=step(lm.tab,direction="backward",k=log(40))
summary(step.tab1) step.tab1$anova
```

C.56 Choix de modèle par la méthode ascendante

```
step.tab2=step(lm.tab,direction="forward",k=log(40))
summary(step.tab2) step.tab2$anova
```

C.57 Choix de modèle par la méthode progressive

```
step.tab3=step(lm.tab,direction="both",k=log(40))
summary(step.tab3) step.tab3$anova
```

C.58 Sélection automatique du modèle

```
# Chargement de la librairie# library(leaps) #Extraction des
variables explicatives# tab1=tab[,2:6] # Recherche des meilleurs
modèles au sens du  # tab.choix1
=leaps(tab1,tab[,"Recettes"],method="Cp",nbest=1) # Résultats#
tab.choix1$Cp plot(tab.choix1$size-1,tab.choix$Cp) #Meilleur
modèle# t=(tab.choix1$Cp==min(tab.choix1$Cp)) # Liste des
variables explicatives # colnames(tab1)[tab.choix1$whi[t]]

# Recherche des meilleurs modèles au sens du  ajusté #
tab.choix2=leaps(tab1,tab[,"Recettes"],method="adjr2",nbest=1)
tab.choix2$adjr2 plot(tab.choix2$size-1,tab.choix2$adjr2) #
Variables explicatives du meilleur modèle au sens du  ajusté #
t=(tab.choix2$adjr2==max(tab.choix2$adjr2)) # variables
explicatives du meilleur modèle au sens du  ajusté
colnames(tab1)[tab.choix2$whi[t]]

# Liste des meilleurs modèles pour chaque dimension # for(i in
(1:5)) {cat(tab.choix2$adjr2[i],tab.choix1$Cp[i],
colnames(tab1)[tab.choix1$whi[i,]],"\n")}
```

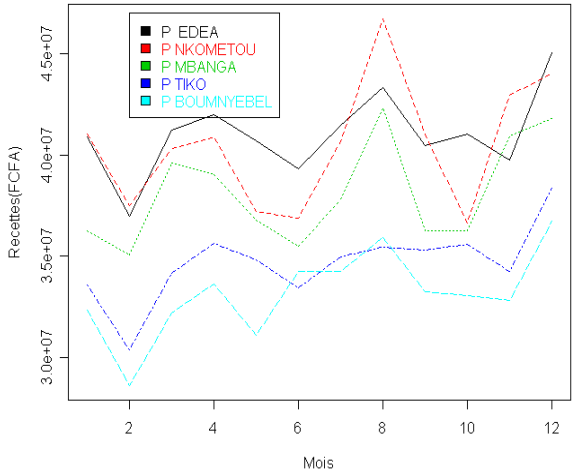
C.59 Evaluation à postériori du modèle

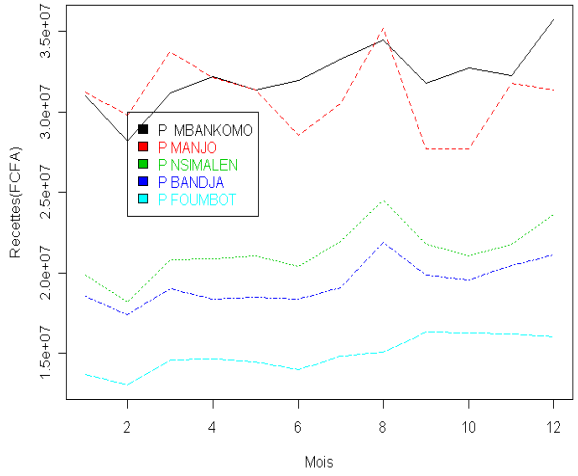
```
plot(tab$Recettes, xlab="années",ylab="Recettes",type="l")
lines(predict(REG),col=2) legend(x=8,y=500000000,c(" Recettes
réelles"," Recettes Prédites"), col=1:2,fill=1:2, text.col=1:2)
title(main=list("FIG.4.12 Courbe des Recettes réelles et des
Recettes prédites"))
```

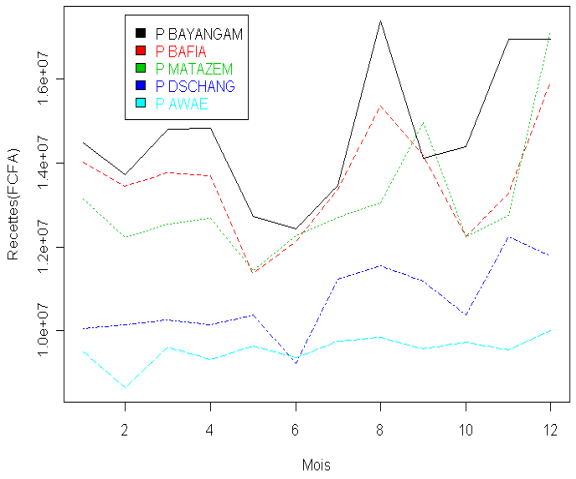
C.60 Erreurs quadratiques d'ajustement

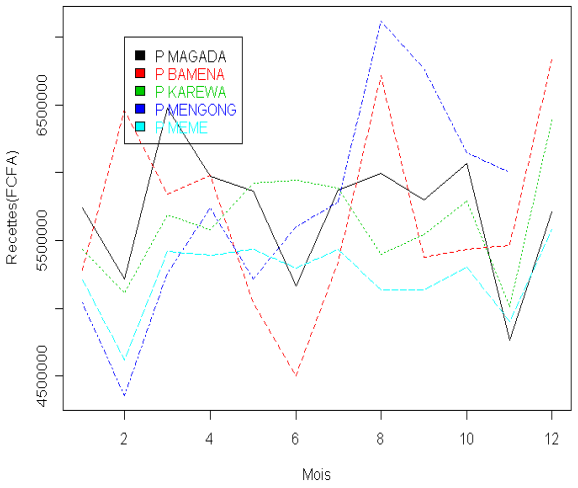
```
# Erreur d'ajustement en prevision générale# r1=reg1$residuals
er1=sqrt(sum(r1^2,na.rm=T)/sum(ts.tab^2,na.rm=T))*100
em1=sqrt(sum(r1^2,na.rm=T)/length(ts.tab)) # Erreur d'ajustement
en prevision corrigée# r2=reg2$residuals
er2=sqrt(sum(r2^2,na.rm=T)/sum(ts.tab^2,na.rm=T))*100
em2=sqrt(sum(r2^2,na.rm=T)/length(ts.tab))
```

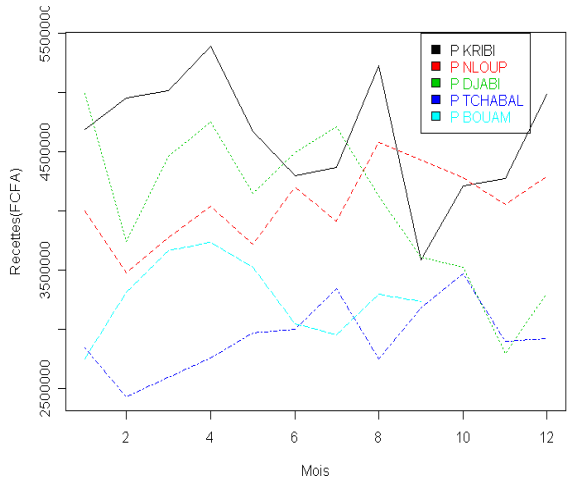
6.2 La présentation des différents tableaux et figures

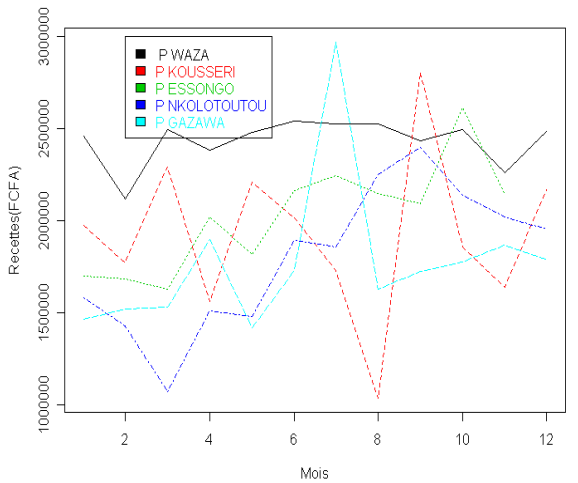


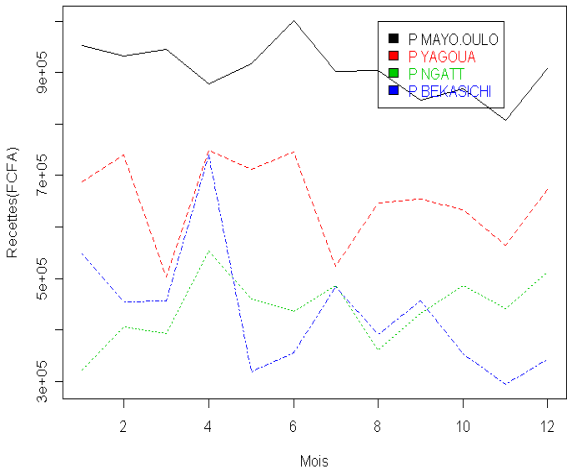


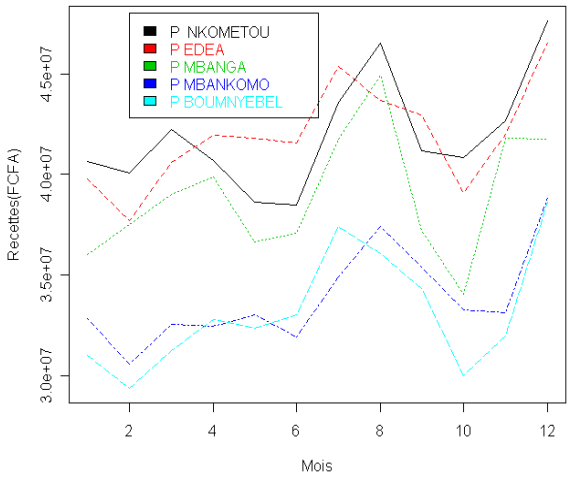


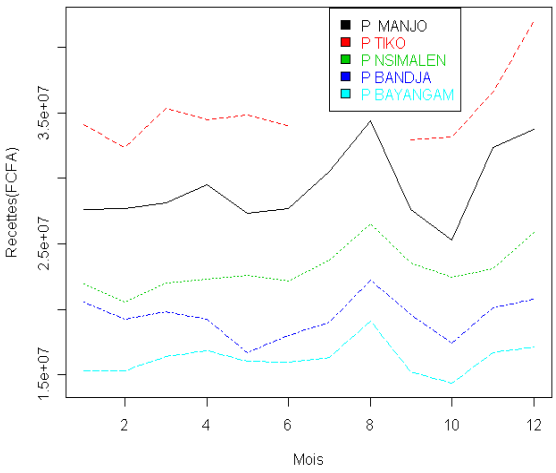


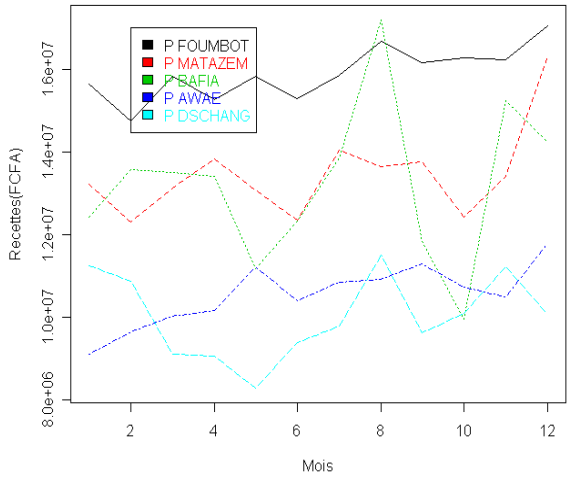


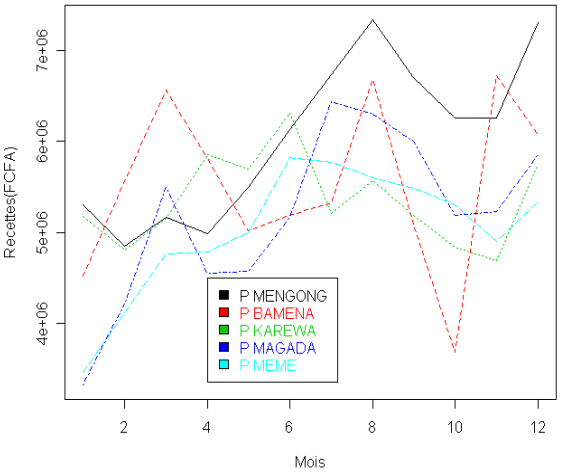


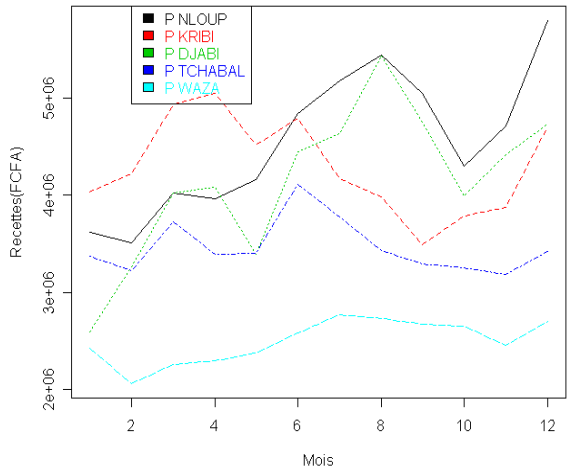

P MAYO.OULO
P YAGOUA
P NGATT
P BEKASICHI
Recettes(FCFA)
3e+05
5e+05
7e+05
9e+05
2
4
6
8
10
12
Mois

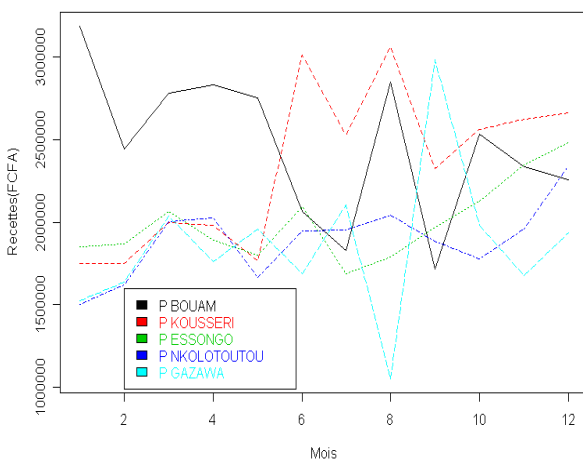

P NKOMETOU
P EDEA
P MBANGA
P MBANKOMO
P BOUMNYEBEL
Recettes(FCFA)
3.0e+07
3.5e+07
4.0e+07
4.5e+07
2
4
6
8
10
12
Mois

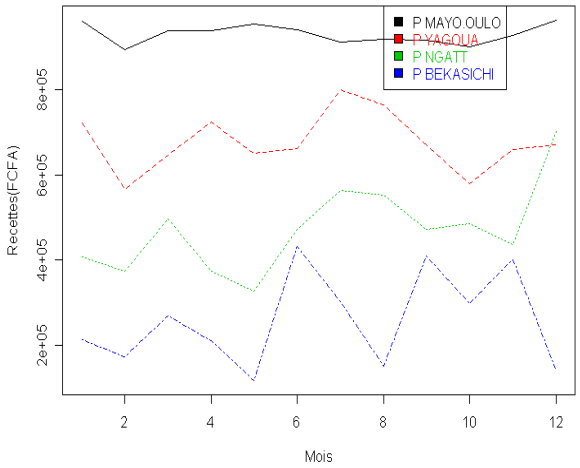

P MANJO
P TIKO
P NSIMALEN
P BANDJA
P BAYANGAM
Recettes(FCFA)
1.5e+07
2.5e+07
3.5e+07
2
4
6
8
10
12
Mois

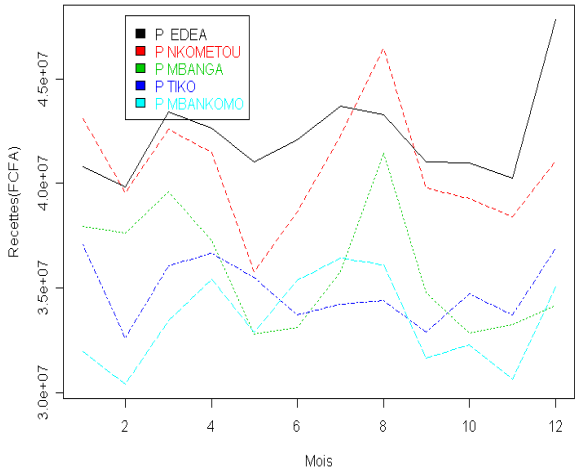

P FOUMBOT
P MATAZEM
P BAFIA
P AWAE
P DSCHANG
Recettes(FCFA)
8.0e+06
1.0e+07
1.2e+07
1.4e+07
1.6e+07
2
4
6
8
10
12
Mois

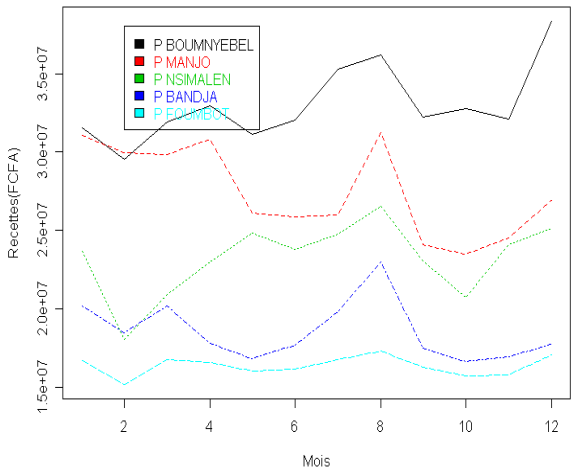

P MENGONG
P BAMENA
P KAREWA
P MAGADA
P MEME
Recettes(FCFA)
4e+06
5e+06
6e+06
7e+06
2
4
6
8
10
12
Mois

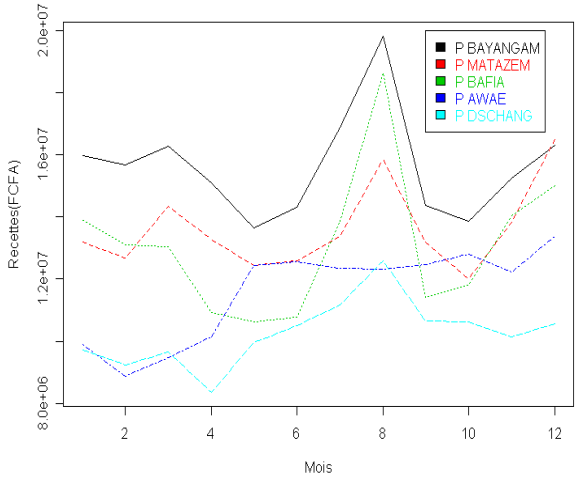

P NLOUP
P KRIBI
P DJABI
P TCHABAL
P WAZA
Recettes(FCFA)
2e+06
3e+06
4e+06
5e+06
2
4
6
8
10
12
Mois

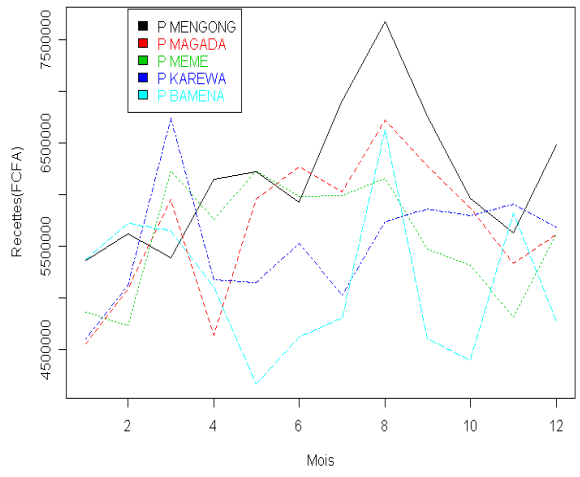

Recettes(FCFA)
1000000
1500000
2000000
2500000
3000000
2
4
6
8
10
12
Mois
P BOUAM
P KOUSSERI
P ESSONGO
P NKOLOTOUTOU
P GAZAWA

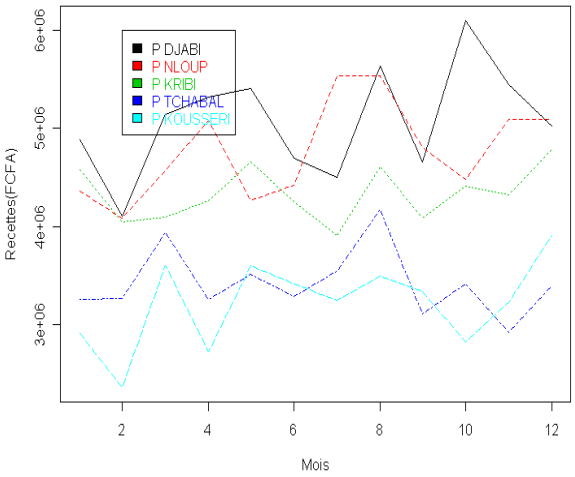

Recettes(FCFA)
2e+05
4e+05
6e+05
8e+05
2
4
6
8
10
12
Mois
P MAYO OULO
P YAGOUA
P NGATT
P BEKASICHI

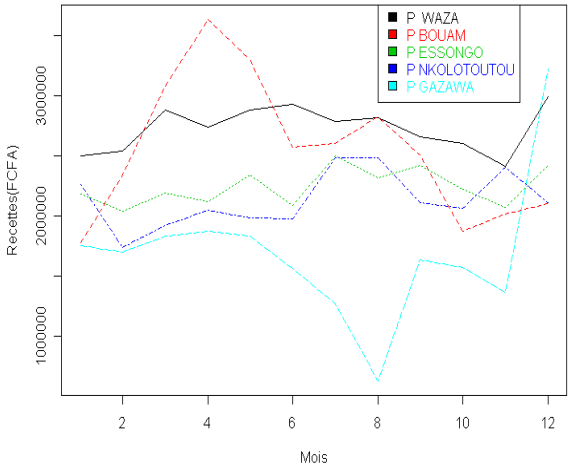

Recettes(FCFA)
3.0e+07
3.5e+07
4.0e+07
4.5e+07
2
4
6
8
10
12
Mois
P EDEA
P NKOMETOU
P MBANGA
P TIKO
P MBANKOMO

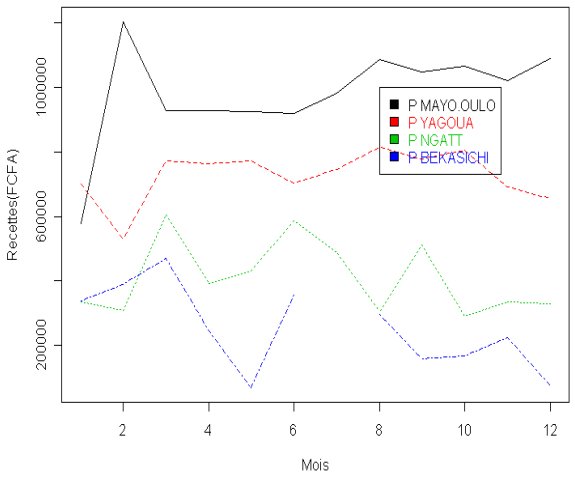

Recettes(FCFA)
1.5e+07
2.0e+07
2.5e+07
3.0e+07
3.5e+07
2
4
6
8
10
12
Mois
P BOUMNYEBEL
P MANJO
P NSIMALEN
P BANDJA
P EDUMBOT

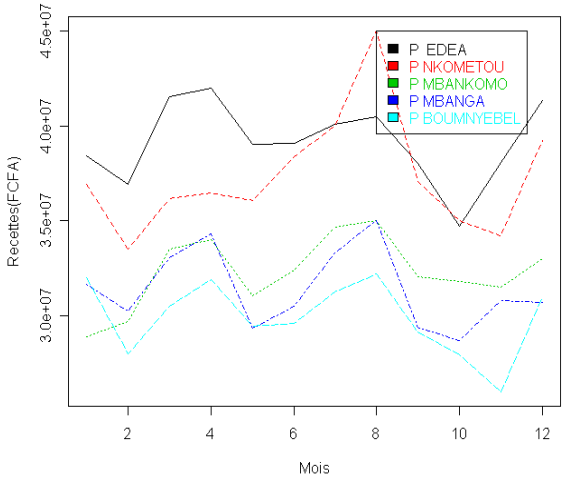

Recettes(FCFA)
8.0e+06
1.2e+07
1.6e+07
2.0e+07
2
4
6
8
10
12
Mois
P BAYANGAM
P MATAZEM
P BAFIA
P AWAE
P DSCHANG

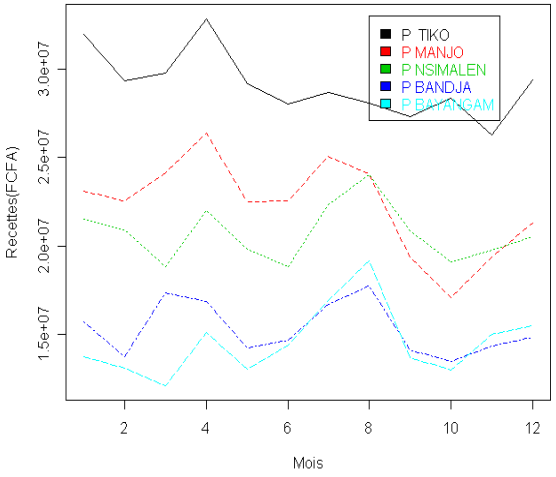

Recettes(FCFA)
4500000
5500000
6500000
7500000
2
4
6
8
10
12
Mois
P MENGONG
P MAGADA
P MEME
P KAREWA
P BAMENA

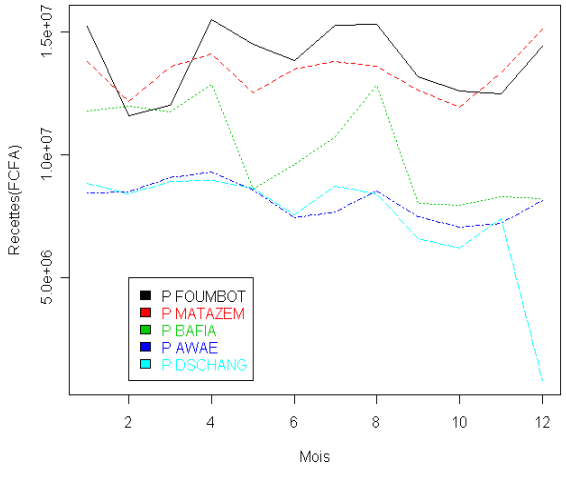

P DJABI
P NLOUP
P KRIBI
P TCHABAL
P KOUSSERI
Recettes(FCFA)
Mois

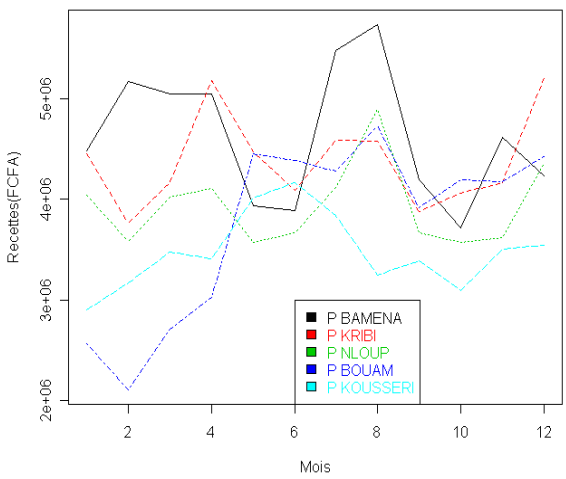

P WAZA
P BOUAM
P ESSONGO
P NKOLOTOUTOU
P GAZAWA
Recettes(FCFA)
Mois

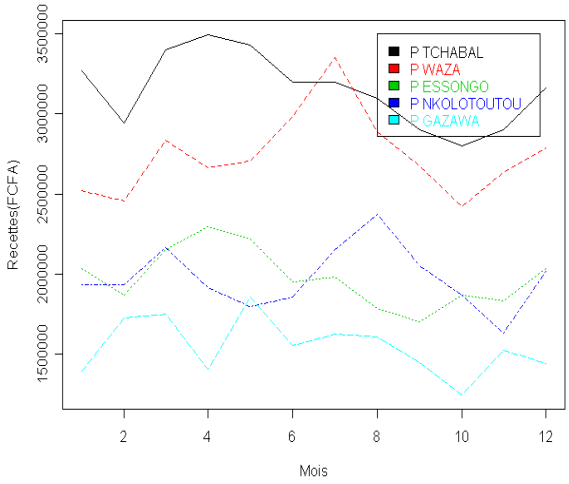

P MAYO OULO
P YAGOUA
P NGATT
P BEKASICHI
Recettes(FCFA)
Mois

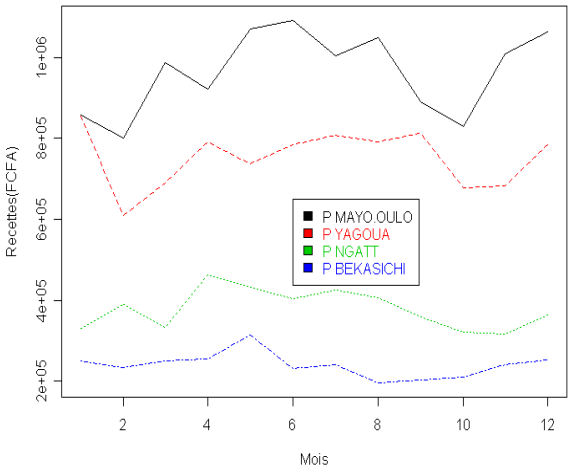

P EDEA
P NKOMETOU
P MBANKOMO
P MBANGA
P BOUMNYEBEL
Recettes(FCFA)
Mois


P TIKO
P MANJO
P NSIMALEN
P BANDJA
P BAFANGAM
Recettes(FCFA)
Mois


P FOUMBOT
P MATAZEM
P BAFIA
P AWAE
P DSCHANG
Recettes(FCFA)
Mois


Recettes(FCFA)
2e+06
3e+06
4e+06
5e+06
2
4
6
8
10
12
Mois
P BAMENA
P KRIBI
P NLOUP
P BOUAM
P KOUSSERI


Recettes(FCFA)
1500000
2000000
2500000
3000000
3500000
2
4
6
8
10
12
Mois
P TCHABAL
P WAZA
P ESSONGO
P NKOLOTOUTOU
P GAZAWA


Recettes(FCFA)
2e+05
4e+05
6e+05
8e+05
1e+06
2
4
6
8
10
12
Mois
P MAYO.OULO
P YAGOUA
P NGATT
P BEKASICHI

Bibliographie

[1] Arthur CHARPENTIER. *Cours de séries temporelles* : DESS Mathématique de la décision et DESS actuariat.

[2] Philippe BESSE. *Choix de modèle en régression linéaire :* Polycopié de cours.

[3] Pascal DKEGNE SIELENOU. *Modélisation et prévision des débits naturels journaliers du B.V.I. à la station de contrôle de SONGMBENGUE :* Mémoire de master 2 de Statistique appliquée, ENSP, Yaoundé I, (2006).

[4] Jean Jacques DROESBEKE, Bernard FICHET, Philippe TASSI. *Théorie et pratique des modèles ARIMA,* Economica.

[5] Olaf KOUAMO . *Prise en compte des données manquantes dans l'analyse d'une étude de cohorte : cas du projet DARVIR :* Mémoire de master 2 de Statistique appliquée, ENSP, Yaoundé I, (2006).

[6] Michel NDOUMBE NKENG. *Modèle linéaire et ses extensions :* cours de Master 2 de Statistiques appliquées, ENSP, Yaoundé I, (2007).

[7] NINO Sylverio. *Séries chronologiques :* Polycopié de cours, (2005).

[8] *Etude de simplification et d'harmonisation de la fiscalité routière au Cameroun.* Ministère des Transports, 1999.

[9] *Etude du système de taxation des véhicules routiers au Cameroun :* Ministère des Travaux Publics et des Transports. LAVALIN INTERNATIONAL, (Mars 1991).

Table des matières

Dédicaces **i**

Remerciements **ii**

Abréviations **iii**

Résumé **iv**

Abstract **v**

Table des figures **vi**

Liste des tableaux **vii**

Introduction **1**

Résumé exécutif **3**

1 Bilan et problématique du peage routier Camerounais **14**
1.1 Historique 14
1.2 Bilan financier 15
1.2.1 Evolution progressive des recettes 15
1.2.2 La fraude au péage 18
1.2.3 l'importance du coût de fonctionnement de péage 18
1.3 Bilan organisationnel et fonctionnel 19
1.3.1 Bilan organisationnel 19
1.3.2 Bilan fonctionnel 21
1.4 Problématique 22

2 Présentation et d'escription des données **24**
2.1 Présentation des données 24
2.2 Description des données 25

3 Méthodes statistiques **30**
3.1 Séries chronologiques 30
3.1.1 Modèles déterministes 30
3.1.2 Modèles stochastiques 34
3.1.3 Vérification à posteriori et choix du modèle 38
3.2 Techniques d'imputation 39
3.2.1 Imputation par la moyenne 40
3.2.2 Utilisation d'un modèle de régression 40
3.2.3 Imputation par une valeur observée tirée au hasard 40
3.3 Régression linéaire multiple 40

4 Applications et résultats. **43**
4.1 Prévisions temporelles 43
4.1.1 Prévision simple : Tendance générale 43
4.1.2 Prévision corrigée : en tenant compte des variations saisonnières 45
4.1.3 Modélisation des recettes mensuelles du péage routier par un processus ARIMA 48
4.2 Imputation des données manquantes 54
4.3 Modélisation et prévision des recettes mensuelles du peage en fonction d'autres paramètres. 55
4.3.1 Choix de modèle " à la main " par élimination 56
4.3.2 Sélection automatique du modèle 57
4.3.3 Dernières estimations 59
4.4 Erreurs quadratiques d'ajustement 60

5 Conclusion générale **62**

Conclusion générale **62**

6 Annexes : Programmes R utilisés **63**
6.1 Commandes et programmes R utilisés 63
6.2 La présentation des différents tableaux et figures 78

Bibliographie **83**

Table des figures

1 Tendance des recettes du peage routier Camerounais. 4
2 Graphe de la série corrigée des variations saisonnières. 8

1.1 Histogramme des recettes du peage Camerounais. 17
1.2 Recettes mensuelles du péage routier suivant les différents exercices budgétaires. . 23

2.1 Evolution des recettes mensuelles du peage routier Camerounais. 26
2.2 Evolution mensuelle des recettes du peage routier Camerounais de 1995 à 1998. . . 26
2.3 Evolution mensuelle des recettes du peage routier Camerounais de 1999 à 2002. . . 27
2.4 Evolution mensuelle des recettes du peage routier Camerounais de 2003 à 2007. . . 27
2.5 Evolution mensuelle du TMI. 28
2.6 Evolution mensuelle du VPAC. 28
2.7 Evolution mensuelle du PSP. 28
2.8 Evolution mensuelle du PGP. 29

4.1 Tendance des recettes du peage routier Camerounais. 45
4.2 Série des moyennes mobiles d'ordre 12. 46
4.3 Graphe de la série corrigée des variations saisonnières. 48
4.4 Evolution des recettes standardisées. 49
4.5 Corrélogramme des recettes standardisées. 50
4.6 Evolution des recettes standardisées et différenciées. 51
4.7 ACF des recettes standardisées et différentiées deux fois. 52
4.8 PACF des recettes standardisées et différentiées. 52
4.9 Diagnostic des résidus ε_t du modèle $ARIMA(5,1,18)$. 55
4.10 Ajustement de la série des résidus. 56
4.11 Diagnostics d'influence des résidus du modèle par élimination. 57
4.12 Diagnostics d'influence des résidus. 60
4.13 Evolution des recettes réelles et des recettes prédictes. 61

Liste des tableaux

1 Prévisions sur les recettes du péage de 2006 à 2008 à base du modèle (1). 5
2 Coefficients saisonniers des différents mois. 6
3 Coefficients saisonniers des différents mois. 7
4 Erreurs quadratiques d'ajustement. 9

1.1 Axes routiers à péage et nombre de postes de contrôle. 16
1.2 Evolution annuelle des recettes du péage routier. 17
1.3 Répartition de la recette du péage. 19

4.1 Prévisions par (4.1) des recettes du péage de 2006 à 2008. 44
4.2 Coefficients saisonniers des différents mois. 46
4.3 Prévisions par (4.2) des recettes du péage de 2006 à 2008. 47
4.4 Paramètres estimés du modèle $ARIMA(5, 1, 22)$. 54
4.5 Résumé des données de notre étude. 56
4.6 Recherche du meilleur modèle au sens du du C_p et du R^2. 58
4.7 Erreurs quadratiques d'ajustement 60